“DID HE REALLY SAY THAT?”
SURVEY EVIDENCE IN DECEPTIVE ADVERTISING LITIGATION

By William W. Vodra and Randall K. Miller

“There’s glory for you!”

“I don’t know what you mean by ‘glory,’” Alice said.

Humpty Dumpty smiled contemptuously.

“Of course you don’t—till I tell you. I meant ‘there’s a nice knock-down argument for you!’”

“But ‘glory’ doesn’t mean ‘a nice knock-down argument,’” Alice objected.

“When I use a word,” Humpty Dumpty said, in a rather scornful tone, “it means just what I choose it to mean—neither more nor less.”1

I. INTRODUCTION

At a time when the Supreme Court has expanded the federal courts’ duty to ensure that proffered scientific and technical testimony is “relevant and reliable,”2 some courts appear to be

* Mr. Vodra is a senior partner and Mr. Miller is a senior associate in the Washington, D.C. law firm of Arnold & Porter, Associate Member of the International Trademark Association. Messrs. Vodra and Miller served as counsel to Pfizer Inc. and UCB Pharma, Inc. in the primary litigation discussed in this article. Although much of the legal and scientific research described in this article was funded by these clients, the views expressed herein are solely those of the authors.

1. Lewis Carroll, The Annotated Alice: Alice’s Adventures In Wonderland and Through The Looking Glass (1960) (with annotations by Martin Gardner) (commenting on this passage, Mr. Gardner writes: “Even in logic and mathematics, where terms are usually more precise than in other subject matters, enormous confusion often results from a failure to realize that words mean ‘neither more nor less’ than what they are intended to mean.”).

retreating from this duty in deceptive advertising cases. Following the Second Circuit’s decision in Schering Corp. v. Pfizer Inc. and UCB Pharma, Inc.,\textsuperscript{3} holding that methodological flaws in surveys properly go to the weight rather than the admissibility of survey evidence,\textsuperscript{4} trial judges are admitting into evidence unscientific surveys,\textsuperscript{5} contrary to the federal courts’ “gatekeeping” obligation.\textsuperscript{6} Worse still, the courts are failing to recognize and apply fundamental tenets of science when evaluating the weight to be afforded admitted surveys, thus accepting inherently unreliable hearsay evidence in deceptive advertising cases.

Schering v. Pfizer involved claims alleging deceptive promotion of Pfizer’s prescription antihistamine ZYRTEC in violation of the Lanham Act.\textsuperscript{7} The litigation spanned approximately five years, produced two complex settlement agreements, and generated several interim judicial decisions, including a significant decision about survey evidence from the Second Circuit in August 1999. This circuit is very influential in

\begin{quote}
fields of expert testimony regardless of the field of expertise (i.e., no longer limited to fields historically considered to be ‘scientific’)).
\end{quote}

\textsuperscript{3} 189 F.3d 218 (2d Cir. 1999).
\textsuperscript{4} See id. at 228.
\textsuperscript{6} See Kumho Tire, 526 U.S. at 141.
\textsuperscript{7} Section 43(a) of the Lanham Act prohibits deceptive advertising; that section provides:

(a)(1) Any person who, on or in connection with any goods or services, or any container for goods, uses in commerce any word, term, name, symbol, or device, or any combination thereof, or any false designation of origin, false or misleading description of fact, or false or misleading representation of fact, which—

* * *

(B) in commercial advertising or promotion, misrepresents the nature, characteristics, qualities, or geographic origin of his or her or another person’s goods, services, or commercial activities, shall be liable in a civil action by any person who believes that he or she is or is likely to be damaged by such act.

the area of Lanham Act litigation, because so many national advertisers, advertising agencies, and media reside in its jurisdiction. Therefore, this decision bears scrutiny.

The basic issue in Schering involved what Pfizer’s field representatives said to physicians in face-to-face sales calls. The parties had previously agreed that certain claims were permitted and others were precluded. At trial, Schering conceded that Pfizer’s printed materials made only permitted claims, but alleged that oral statements made by Pfizer’s sales representatives were prohibited by Section 43(a) of the Lanham Act and the parties’ agreement. To prove what was actually said, Schering relied exclusively on survey evidence from unidentified physicians who had recently been visited by Pfizer representatives.

Complicating the situation further, Schering asserted that the prohibited claim allegedly made—that ZYRTEC was “nonsedating”—involved a particular meaning of the term “nonsedating.” Pfizer accepted that, by Schering’s narrow definition, the claim was false. But Pfizer argued that the survey evidence could not establish that physicians, a “sophisticated” consumer audience, used the term “nonsedating” when responding to survey questions in the rigid way Schering contended. Like Alice, Pfizer noted that a word does not always mean what one person wants it to mean.

This article contends that the Second Circuit accepted Schering’s assumptions in lieu of demanding reliable scientific proof by ordering the lower court to admit the surveys into evidence under exceptions to the hearsay rule. On remand, the trial court further gave the surveys sufficient weight to issue a preliminary injunction against Pfizer. Subsequently, the litigation was resolved by settlement.

Often, legal scholars find it difficult to interpret decisions such as this, which grant interim relief on a hastily assembled and incomplete record, and where settlement precludes further development and explication of the factual and legal principles. In this case, however, while preparing for trial on the merits, Pfizer commissioned scientific research to test its arguments against the Schering surveys, that is, that the term “nonsedating” in physician responses to surveys did not reliably show that Pfizer sales reps used that term in its promotional activities. This research has now been published.8 It established that the survey methods on which the circuit and district courts relied were unreliable—and produced unreliable results. This article presents information to emphasize why courts must understand and require appropriate

---

scientific standards for surveys before they are admitted into
evidence.

Part II of this article describes the background and judicial
decisions from the Schering v. Pfizer litigation. Part III discusses
recently published survey experiments that expose the fallacy of
key assumptions made by the Second Circuit in Schering. Part IV
examines a plaintiff’s legal and scientific burden in Lanham Act
deception cases and explains that surveys relied upon to establish
the causal link between the defendant’s conduct and competitive
injury must be evaluated as scientific experiments requiring
proper experimental designs, such as controls, in order to be
considered relevant and reliable.

II. SCHERING v. PFIZER:
A CASE STUDY IN BAD SCIENCE
MAKING BAD LAW

The Schering v. Pfizer deceptive advertising litigation involved
competing prescription antihistamines and competing marketing
claims related to a single side effect—drowsiness—associated with
many antihistamine products. As with all litigation involving
allegations of deceptive advertising or unfair competition, it is
necessary to understand the relevant marketplace in order to
study the case. But this case, like virtually all Lanham Act
litigation, also was affected by the rapid pace and limited
evidentiary record before the courts. The legal conclusions were
influenced by the fact that “good science” was not then available to
demonstrate the flaws in the evidence presented. Accordingly, an
examination of this case should begin with its procedural history.

A. Procedural Posture and Impact in Litigation

The 1998 litigation was typical of “emergency” injunction
litigation under the Lanham Act that proceeds on an expedited
schedule, often at a breakneck pace. The purpose of this rapid-
pace procedure is to ensure that a plaintiff legitimately aggrieved
can halt improper advertising before it inflicts irreparable harm,
that is, harm such as impairment of good will, business reputation,

where a plaintiff can demonstrate an urgent need for immediate judicial intervention to
avoid a specific type of injury for which there is no adequate remedy at law. “[P]reliminary
injunctions are generally granted under the theory that there is an urgent need for speedy
action to protect the plaintiff’s rights.” Synagro-WWT, Inc. v. Louisa County, No. Civ. A.
Citytrust, 756 F.2d 273, 276 (2d Cir. 1985)). The reverse of this principle is that even a brief
delay between the discovery of alleged wrongdoing and initiation of an action at equity “may
be construed as inconsistent with the sense of necessity ordinarily associated with the need
for such a remedy.” Id.
and diversion of market share, which is difficult to quantify in terms of monetary damages. Irreparable harm is of particular concern in cases where a deceptive advertising claim specifically mentions a competitor or a competitor's product by name; the reputation of a business or product, once maligned with persuasive yet deceptive advertising, is difficult to restore. Thus, the benefit of accelerated procedures is the opportunity to invoke federal equity power to halt deceptive advertising before it causes injuries for which remedies at law simply are inadequate.

The abbreviated expedited procedure permits, if warranted, issuance of interim injunctive relief as a temporary, stopgap measure to preserve the status quo until the court can entertain a full and final hearing on the merits. This expedited procedure, however, creates disadvantages for both parties and courts. The parties do not have the benefit of a complete discovery period to develop evidence to be used at the preliminary injunction hearing. Courts attempt to ameliorate this problem by ordering very limited, narrowly targeted discovery on an accelerated schedule to allow the parties at least some preparation for the preliminary injunction hearing. Nevertheless, the courts are deprived of a fully developed evidentiary record; thus, decisions must be reached before all useful information can be presented. This leads to a dilemma in deceptive advertising cases in that the plaintiff bears the burden to demonstrate the reliability of survey

10. See Tom Doherty Assocs., Inc. v. Saban Entm't, Inc., 60 F.3d 27, 37-38 (2d Cir. 1995) (irreparable harm is "very difficult to quantify at trial" and "for which a monetary award cannot be adequate compensation"). Irreparable harm, however, cannot be speculative or remote. See Coca-Cola Co. v. Tropicana Prods., Inc., 690 F.2d 312, 316 (2d Cir. 1982) ("The Lanham Act plaintiff must . . . offer something more than a mere subjective belief that he is likely to be injured . . . he must submit proof.").

11. In cases where deceptive advertising mentions a competitor or competing product by name, irreparable harm is presumed. See, e.g., Cashmere & Camel Hair Mfrs. Institute v. Saks Fifth Ave., 284 F.3d 302, 314-15 (1st Cir. 2002); S.C. Johnson & Son, Inc. v. Clorox Co., 241 F.3d 232, 238 (2d Cir. 2001).

12. See, e.g., Licata & Co., Inc. v. Goldberg, 812 F. Supp. 403, 407 (S.D.N.Y. 1993) ("Where events with irrevocable effects which are difficult to unscramble are imminent, quick resolution on the merits is crucial to those involved"). Cf. Castrol, Inc. v. Pennzoil Co., 799 F. Supp. 424, 440 (D.N.J. 1992) (trademark case) ("Whenever competitive unfairness exists, whether it be of . . . a disrespect for intellectual property, or a false advertising claim, swift and certain penalties must attach to the offending conduct."). aff'd, 987 F.2d 939 (3d Cir. 1993). The rapid pace, however, underscores the need to hold a plaintiff to its burden. "Quick lawsuits with too easy a burden of proof against small competitors could have a stifling effect on competition." Id. at 951 n.1.

13. See, e.g., Thomas & Betts Corp. v. Panduit Corp., 138 F.3d 277, 292 (7th Cir. 1998) (noting the preliminary injunction decisions "are often based on incomplete evidence and a hurried consideration of the issues").

but also wants expedited consideration on a schedule that may not permit sufficient time for the defendants or the courts to evaluate the reliability of survey design.

The 1998 Schering v. Pfizer litigation illustrates these disadvantages. Schering initiated it with a motion for a temporary restraining order (TRO) on October 5, 1998. As discussed in more detail below, Schering presented survey data in support of the motion suggesting that Pfizer was promoting ZYRTEC as “nonsedating.” The district court heard oral argument on Schering’s TRO motion two days later. The court denied Schering’s motion, in part based upon the parties’ willingness to schedule a preliminary injunction hearing quickly to commence three weeks later. The court also noted its reluctance to issue a TRO based on survey evidence when the defendant had no opportunity to take “relevant discovery regarding the preparation of the survey” or “to cross examine” the survey proponents.

In the interval between the denial of Schering’s TRO motion on October 7, 1998, and the commencement of the preliminary injunction hearing, the parties undertook expedited discovery, which compressed much activity into a short period of time. During the preliminary injunction hearing, Pfizer presented a survey expert, Dr. Ivan Ross, to testify about methodological weaknesses in Schering’s survey data; however, there was insufficient time to prepare new studies to demonstrate how fatal these flaws were to Schering’s surveys.

After the three-day preliminary injunction hearing, neither party undertook further discovery or new surveys, but awaited a ruling on the preliminary injunction motion. Notwithstanding the expedited procedure, however, the court did not issue its decision until March 1999, when it denied Schering’s motion for preliminary injunction.

The United States Court of Appeals for the Second Circuit granted Schering’s motion for an expedited appeal, and additional discovery and new studies or surveys were deferred pending the outcome of the appeal. On August 17, 1999, the Second Circuit

---


18. Id. The district court added: “I have had in my experience very good and very useful surveys and I have had some stinkers.” Id.

vacated the district court’s denial of Schering’s preliminary injunction motion and remanded for reconsideration.20

On remand, by agreement of the parties, the district court entertained one afternoon of oral argument but did not receive any additional evidence. In other words, the district court again had before it the same evidentiary record that was hurriedly assembled in the three-week interval between the denial of the TRO and the three-day preliminary injunction hearing, an evidentiary record that was now over a year old.

Following this argument, again, neither party initiated further discovery or new research. In May 2000, the district court issued an opinion granting in part and denying in part Schering’s motion for preliminary injunction.21

At this point, Pfizer began preparation for a full trial on the merits of Schering’s pending motion for a permanent injunction. As part of its activities, Pfizer commissioned several scientific experiments to test the validity of Schering’s survey evidence. For various reasons, the parties decided to settle in April 2001. As a result, these investigations were never considered by any court in this matter. One purpose of this article is to put this research into the public domain, with the hope it may benefit courts and parties in future litigation.

B. ZYRTEC, CLARITIN, and the Antihistamine Market

1. Antihistamines

Antihistamines are drugs intended to relieve symptoms such as runny nose, watery eyes, postnasal drip, and skin rash and itching that occur following the body’s release of histamine triggered by allergic reactions to things like pollen, dust mites, and animal dander.22 The antihistamine market includes both drug products that can only be obtained with a physician’s prescription and those that can be purchased at retail stores over the counter (OTC) without a prescription. Familiar OTC antihistamines include BENADRYL (diphenhydramine), CHLOR-TRIMETON (chlorpheniramine), DRIXORAL (dexbrompheniramine), and TAVIST (clemastine fumarate). Such OTC antihistamines have been on the market for decades and are frequently referred to as “first generation” antihistamines. In the late 1980s, a new group of products was introduced, which became known as “second generation” antihistamines and which required a prescription. The

20. Schering Corp. v. Pfizer Inc. and UCB Pharma, Inc., 189 F.3d 218 (2d Cir. 1999).


early members of this new group were SELDANE (terfenadine) and HISMANAL (astemizole), both of which have been removed from the market for safety reasons. Later members are CLARITIN (loratadine), ZYRTEC (cetirizine), and ALLEGRA (fexofenadine).

A significant difference between the first and second generation antihistamines lies in their propensity to cause drowsiness (also referred to by the Food and Drug Administration (FDA) as “somnolence”), an undesirable side effect in most circumstances. Because first generation antihistamines cause a high percentage of users to experience drowsiness, FDA requires that their labels carry a “warning” about possible drowsiness.

The second generation prescription antihistamines generally cause a smaller percentage of users to experience drowsiness, and when drowsiness is experienced, it is generally less severe. For some products, such as ALLEGRA (made by Aventis) and Schering’s CLARITIN, the proportion of patients who experience this side effect at recommended doses is not statistically different from placebo. In others, such as Pfizer’s ZYRTEC, the rate is statistically higher for the drug than for placebo.

FDA requires that the labeling for both CLARITIN and ZYRTEC disclose the incidence of drowsiness so that physicians can make informed clinical judgments when prescribing these products to their patients. The incidence rates of drowsiness—the percentage of the population that experienced the drowsiness side effect in clinical trials—reported in the FDA-approved labeling for ZYRTEC and CLARITIN are as follows:

---

25. See Jean Bousquet et al., H1-Receptor Antagonists: Structure and Classification in Histamine and H1-Receptor Antagonists in Allergic Disease 91, 101-02 (F. Estelle R. Simons ed., 1996); Jacquelynne P. Corey, Advances in the Pharmacotherapy of Allergic Rhinitis: Second—Generation H1 Receptor Antagonists, 109 Otolaryngology Head Neck Surgery 584, 591 (Sept. 1993) (“[I]mportant improvements in second generation H1 antagonists have been made compared with first generation antihistamines because of the marked decrease in sedation potential.”).
26. 21 C.F.R. § 341.72(c)(3), (4). Although more patients experience somnolence with first generation antihistamines compared to second generation antihistamines, not all patients will experience the side effect. See Warner-Lambert Co. v. Schering-Plough Corp., No. 91 Civ. 5079, 1991 U.S. Dist. LEXIS 14620 (S.D.N.Y. Oct. 15, 1991) (permanently enjoining Schering from airing a television commercial that falsely suggested that all or almost all users of BENADRYL experience drowsiness).
<table>
<thead>
<tr>
<th>Adults</th>
<th>ZYRTEC</th>
<th>14% vs. 6% placebo</th>
<th>CLARITIN</th>
<th>8% vs. 6% placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Children</td>
<td>ZYRTEC</td>
<td>4.2% vs. 1.3% placebo</td>
<td>CLARITIN</td>
<td>2% vs. 4% placebo</td>
</tr>
</tbody>
</table>

The dizziness incidence rates for ZYRTEC mean that in the clinical trials for ZYRTEC, approximately 86 percent of adults (100-14=86 percent) and 95 percent of children (100-4.2=95.8 percent) reported experiencing no dizziness after taking ZYRTEC.28 The labeling information for ZYRTEC also states that: (1) ZYRTEC’s side effects (including dizziness) are “mild or moderate”; and (2) Discontinuations due to somnolence are uncommon (1.0 percent ZYRTEC vs. 0.6 percent placebo).

Both the benefits of antihistamines and side effects may vary with the size of the dose. This “dose-response” relationship means that if one exceeds the “recommended” dose or doses, the likelihood of side effects may increase.29 For example, in clinical trials, the incidence of somnolence seen with 10 mg CLARITIN tablets “was similar to that seen with placebo,” but when patients took higher doses of CLARITIN, somnolence occurred more frequently than with placebo. Nevertheless, because the somnolence caused by CLARITIN’s recommended dose of 10 mg is not significantly different from the somnolence observed with patients taking a placebo, FDA permits CLARITIN to be advertised as a “nonsedating” product.30 Such a marketing claim may not be made for ZYRTEC in the United States because, even though ZYRTEC causes far fewer patients to experience somnolence when compared to first generation antihistamines, the number of patients who experience somnolence with 10 mg ZYRTEC tablets (14 percent) is statistically higher than with patients who take a placebo (6 percent).

In sum, the somnolence associated with different antihistamines—first and second generation—varies with the drug and with the dose recommended. The medical community and regulatory authorities recognize variation along a spectrum. But

30. Oct. 12, 1995 letter to Richard N. Spivey, Schering Corp., from Joan Hankin, Reg. Review Officer, FDA (“The determination of the appropriateness of the use of the term ‘sedating’ versus ‘nonsedating’ is generally based on the adverse event profile (such as somnolence) versus placebo that would be generated during clinical trials. Nonsedating antihistamines would need to have a sedation rate comparable to or preferably less than that observed for placebo.”).
some marketers, such as Schering for its CLARITIN product, can make claims that are proscribed by FDA for other marketers, such as Pfizer’s ZYRTEC product.

2. Competitive Marketing Positions

Schering’s competitive strategy was played against this background. In the marketplace—and in the courts—Schering argued that antihistamines either are “nonsedating” or “sedating.” For Pfizer, however, the issue was part of a spectrum in which the risk of the side effect varied among patients and could be acceptable to physicians and patients because of ZYRTEC’s other benefits.

FDA does not use the term “sedating” or “nonsedating” in prescription drug labeling to describe the drowsiness side effect of antihistamines. Nor does FDA classify antihistamines as “sedating” or “nonsedating.” Notably, FDA has stated that any claim suggesting that antihistamines are categorized by FDA as either “nonsedating” or “sedating” would be deceptive.

Nevertheless, the market positioning urged by Schering remained dichotomous: either an antihistamine was “nonsedating” (as CLARITIN could claim in the United States) or it must be “sedating.” The outcome of the Schering v. Pfizer litigation would be influenced by whether the courts believed that physicians shared this dichotomous view.

3. 1996 Litigation and Its Effect on Competitive Claims

The legal landscape for the 1998-2001 litigation was established to a large degree by a previous round of deceptive advertising litigation between Schering and Pfizer in 1996. Within a few weeks after Pfizer introduced ZYRTEC into the United States market in 1996, Schering filed a deceptive advertising suit under the Lanham Act claiming that Pfizer was falsely promoting ZYRTEC as “nonsedating.” Pfizer counterclaimed, alleging that Schering was falsely exaggerating the incidence of somnolence associated with ZYRTEC. The parties quickly settled.

---

31. Sept. 24, 1996 letter to Ronald J. Garutti, Schering Corp., from Joan Hankin, Reg. Review Officer, FDA (“Contrary to Schering’s assertions, FDA has not developed any such classification scheme to categorize antihistamines as sedating versus nonsedating drugs. Therefore, any suggestion, statement, or representation that FDA has classified antihistamine drugs into categories of sedating or nonsedating would be false and/or misleading.”).


33. The settlement agreement between Pfizer and Schering was executed on April 4, 1996 [hereinafter referred to as the “1996 Settlement Agreement”]; it was made public by
the settlement, among other things, Schering agreed not to claim (expressly or by implication) that ZYRTEC is comparable in incidence or severity of somnolence to “first generation” antihistamines.34 Pfizer agreed not to claim (expressly or by implication) that ZYRTEC is “nonsedating.”35 Both parties retained the right to disseminate (1) accurate factual information, including FDA-approved statements about drowsiness incidence rates and severity; (2) peer-reviewed medical literature; and (3) truthful nonmisleading claims based upon valid clinical data.36 Violations of the 1996 Settlement Agreement would be actionable as breach of contract, as well as (arguably) under the Lanham Act.

The parties’ 1996 litigation and settlement establishes some of the key parameters of the later litigation in 1998 and also provides a context for understanding the prescription antihistamines market. For example, although Schering was prohibited from “lumping” ZYRTEC with first generation antihistamines in terms of the degree of somnolence, Schering arguably was not barred from stating that ZYRTEC is “sedating.” Thus, in 1998 Schering disseminated a promotional piece emblazoned with the advertising slogan, “Zyrtec is a sedating antihistamine! And Pfizer can’t deny it!”37 This and other marketing materials reinforced Schering’s preferred dichotomy that antihistamines were either sedating or nonsedating and that ZYRTEC should be considered part of the former category. Because Pfizer could not claim that ZYRTEC was “nonsedating,” the marketer was challenged to respond effectively with truthful and non-misleading facts.

The battle between Schering (the market leader with roughly 50 percent of the prescription antihistamine market) and Pfizer (with about 17 percent)38 erupted in more litigation in October 1998.

---

34. 1996 Settlement Agreement ¶ 4. Schering also agreed not to claim expressly or by implication that the severity of somnolence or performance impairment with ZYRTEC is comparable to legal intoxication or that a health care professional may be subject to legal liability as a result of prescribing ZYRTEC. Id.

35. 1996 Settlement Agreement ¶ 2; Pfizer also agreed not to expressly state that ZYRTEC was “low sedating.” Id. See also Schering, 2000 WL 718449, at *2 (quoting from the Settlement Agreement).


38. See, e.g., Elyse Tanouye, Allergy Drugs Wage a Bitter War of the Noses, Wall St. J., May 23, 1996, at B1; cf. Dennis Cauchon, Why Allergy Drugs Cost So Much, USA Today, April 12, 2000, at 1A (noting that CLARITIN “is the ninth-best-selling drug in the United States and the most heavily advertised.”).
C. The 1998 Schering v. Pfizer Litigation

The 1998 suit was based entirely on Schering’s allegation that Pfizer was promoting ZYRTEC as “nonsedating.”39 Schering acknowledged that ZYRTEC’s labeling, as well as all printed and audio-visual advertising and promotional literature, accurately summarized and disclosed the fact that somnolence is ZYRTEC’s most common side effect. The problem, Schering claimed, was what Pfizer field representatives were saying in oral communications to physicians in “detailing” sessions. Detailing is a promotional medium commonly employed in the pharmaceutical industry that usually takes the form of one-on-one informational meetings of 5-10 minute duration between a sales representative (or “detail person”) and a physician in the latter’s office.40 Specifically, Schering alleged that Pfizer’s detail persons were telling doctors that ZYRTEC did not cause somnolence, was “nonsedating” or “essentially non-sedating,” or otherwise was free from this side effect.

Schering based its case on market research surveys of physicians who had met with a Pfizer sales representative within the prior one to five days. In various ways, the surveys asked physicians to recall the “main message” of the sales presentation and to remember what, if anything, the sales representative said about ZYRTEC and “sedation.” Schering did not call any individual physician to testify about what the physician was told. Schering also did not present testimony of any current or former Pfizer sales representative to say they had told physicians that ZYRTEC was “nonsedating.” To prove that these communications occurred, Schering offered only survey evidence.41

Pfizer’s response was a flat denial. In the usual Lanham Act suit, the advertiser acknowledges that a marketing claim was made, but disputes that the claim was false or misleading. Here, however, Pfizer acknowledged that the claim, if made, would have been false, but steadfastly asserted the claim was not made. It presented, for example, testimony of persons involved in product marketing, sales training, and actual field sales work, who

39. Complaint, Schering Corp. v. Pfizer Inc. and UCB Pharma, Inc., 98 Civ. 7000 ¶¶ 18, 22 (S.D.N.Y. Oct. 5, 1998) (alleging that Pfizer was engaged in a “systematic and concerted” effort to have its detail representatives tell physicians that ZYRTEC “is non-sedating or essentially non-sedating”).

40. See Schering, 189 F.3d at 222; see also Schering, 2000 WL 718449, at *1.

41. Schering offered the survey to show the central element of its Lanham Act claim, that is, that Pfizer had communicated a false or misleading message. Survey evidence has also been used to support other elements, such as materiality and damages. See, e.g., JTH Tax, Inc. v. H & R Block Eastern Tax Services, Inc., 2002 WL 27257 (4th Cir. Jan. 10, 2002) (survey used to show that an advertising claim was “material” to consumers); Harolds Stores, Inc. v. Dillard Dept Stores, Inc., 82 F.3d 1533 (10th Cir. 1996) (survey used to show impact of a stipulated copyright violation on the plaintiff’s public reputation).
explained that detail persons were trained and routinely monitored as to permissible claims.

Thus, the outcome of the litigation would entirely rest on the admissibility of, and weight given to, Schering’s surveys of physicians’ recollections. Without them, Schering could not sustain its burden of proof.42

The Schering v. Pfizer case requires an appreciation of the difference between two categories of false advertising claims (literal and implied claims) and the application of these two branches of false advertising law to the advertising medium of in-person verbal promotion (as opposed to advertising media that is more traditionally the subject of Lanham Act litigation, such as print, television, or radio). Generally, Lanham Act plaintiffs are obligated to show that a challenged advertising statement falls into one of two primary categories of prohibited commercial speech: either the advertisement is literally false—or the advertisement is literally truthful but has a tendency to deceive a substantial segment of the target audience. When an advertising statement is literally false, a plaintiff is entitled to injunctive relief without any showing, by survey evidence or otherwise, of actual customer deception.45 In implied falsehood cases, the plaintiff must show a significant adverse impact on the target audience; such showing usually requires the introduction of scientifically sound survey evidence.46

42. See Schering Corp. v. Pfizer Inc., 98 Civ. 7000 (LMM), 1999 WL 144921, at *7 (S.D.N.Y. Mar. 16, 1999) (“The Court has considered the other evidence and arguments of Schering in support of its motion for a preliminary injunction. Whatever their weight might be in conjunction with admissible evidence of what the Pfizer and UCB representatives said to the physicians, they are, absent such evidence, an insufficient basis on which to grant a preliminary injunction.”), vacated, 189 F.3d 218 (2d Cir. 1999).

43. Plaintiffs alleging deceptive advertising under Section 43(a) of the Lanham Act must prove that the defendant used a “false or misleading” description or representation of fact in the context of interstate commercial advertising that “misrepresents the nature, characteristics, qualities, or geographic origin” of its own or another’s product. 15 U.S.C. § 1125(a) (emphasis added). An advertisement may be literally false either on its face or based on the message it conveys by “necessary implication.” See, e.g., Castrol Inc. v. Pennzoil Co., 799 F. Supp. 424, 437 (D.N.J. 1992), aff’d, 987 F.2d 939 (3d Cir. 1993).

44. What percentage of consumers qualifies as “substantial” has not been quantified. See Mead Johnson & Co. v. Abbott Labs., 41 F. Supp. 2d 879, 888 n.4 (S.D. Ind. 1999) (collecting cases that discuss the issue, including Johnson & Johnson v. Merck Consumer Pharm. Co. v. Rhone-Poulenc Rorer Pharm., Inc., 19 F.3d 125, 133-34 (3d Cir. 1994), which suggested that 7.5 percent of consumers was not sufficient to prove deception), rev’d on other grounds, 201 F.3d 883 (7th Cir. 2000).

Schering was breaking new ground by using surveys to support a literal falsehood claim, not to prove that the claims were false but to prove the claims were made. Pfizer conceded that the alleged statement—"Zyrtec is non-sedating"—if actually made, was literally false. Schering was attempting to use the survey to show that Pfizer sales representatives spoke certain words that would constitute literally false advertising.

By the time the case reached the Second Circuit, however, Schering had broadened its case and was now arguing that the surveys could be used to show both literal and implied falsehoods.47 As discussed below, this new theory was a key point of confusion in the courts’ opinions. For example, the Second Circuit discussed both branches of false advertising law, and emphasized Schering’s implied falsity claim, whereas the district court on remand held that Schering had not actually litigated an implied claim and thus relied only on the literal falsehood claim.

1. The Surveys Before the Courts

At the preliminary injunction hearing, Schering offered five surveys and testimony of a survey expert, John Bartolomeo.48 Schering’s principal survey was one that was specifically designed for the litigation by Mr. Bartolomeo (as opposed to “off-the-shelf” surveys designed to collect general marketing information).49 The survey began by telephonically contacting physicians until 200 physicians were identified who reported that they had been visited by a Zyrtec representative within the prior 24-48 hours. Of these 200, the survey staff asked these questions:

**Question 1**: During (TODAY’S/YESTERDAY’S) (ADJUST ACCORDING TO ANSWERS IN Q.B2) ZYRTEC detail, what did the rep who did the detail say to you about ZYRTEC tablets? Please be as specific and as complete as possible. (PROBE:) Anything else? (RECORD VERBATIM)

**Question 2**: You may have mentioned this already, but I just want to make sure I have this right. During (TODAY’S/YESTERDAY’S) (ADJUST ACCORDING TO ANSWERS IN Q.B2) ZYRTEC detail,
did the rep say anything about ZYRTEC tablets and sedation?

**Question 3:** Even at the risk of repeating what you have already told me, please be as complete and specific as possible about what was said by the rep about ZYRTEC tablets and sedation during (TODAYS/YESTERDAYS) (ADJUST ACCORDING TO ANSWERS IN Q.B2) ZYRTEC detail. (PROBE:) Anything else? (RECORD VERBATIM)

The survey was conducted between July 23 and September 17, 1998.\(^50\) Based on these “verbatim” answers—quoting what the physician told the interviewer—Mr. Bartolomeo and counsel for Schering\(^51\) together “coded” the answers into common terms for quantification and analysis. According to Schering, 14.5 percent of the physician responses, as coded, indicated that the physician remembered the Pfizer sales representative stating that ZYRTEC is “nonsedating.”

Schering also commissioned a study in which similar questions were asked of a smaller pool of 98 physicians recently (within the last two days) detailed by a ZYRTEC sales representative. The survey firm DTW conducted the survey in March and April 1998. Among other things, the survey asked doctors the following questions:

1. What did the sales representative tell you about the product? Please be as specific and complete as you can in describing the message or information as conveyed to you by the sales representative on the product.

2. What did the representative say about the product and sedation?

Schering’s survey expert and counsel again coded the responses and concluded that 16.4 percent of the physicians indicated that the sales representative said, in substance, that ZYRTEC is “nonsedating.”

Schering also relied on three other sets of “off-the-shelf” market research studies. These studies were neither crafted for the litigation nor designed with the rigor normally demanded of a “litigation survey.”\(^52\) The market research surveys posed a broad

---

\(^{50}\) Schering, 189 F.3d at 223.


\(^{52}\) Off-the-shelf surveys may not be designed to answer the precise litigation question. Cf. Loctite Corp. v. Nat’l Starch & Chem. Corp., 516 F. Supp. 190, 206 (S.D.N.Y. 1981) (marketing surveys conducted before litigation were designed to test for brand awareness, whereas the "single issue at hand . . . [was] whether consumers understood the term 'Super Glue' to designate glue from a single source").
range of questions about a variety of products, but also included specific questions about ZYRTEC detail visits. One such survey, which was also conducted by the DTW firm between March 30 and June 5, asked 78 physicians: “What did the sales representative tell you about ZYRTEC? Please be as specific and complete as you can be in describing the message or information that was conveyed to you about ZYRTEC.” Approximately 14.1 percent of the surveyed physicians provided a response that Schering interpreted as suggesting that the sales agent was asserting that ZYRTEC is “nonsedating.”

A second off-the-shelf survey was conducted by Market Measures, Inc. (“MMI”) and asked 74 physicians:

(1) In one or two sentences, what was the main message of the [detail] presentation?

(2) In one or two sentences, please describe what else the representative discussed.

Schering argued that 20.3 percent of the responses to the MMI survey reflected a claim that ZYRTEC did not cause sedation.

Finally, a survey conducted by Scott Levin collected responses from 716 physicians from March through August 1998 who were asked to identify the main messages being conveyed by sales representatives during ZYRTEC detail sessions. In this case, physicians were not asked what was said, but what they perceived to be the “take-away” message. The following table summarizes the top six messages from ZYRTEC details:

<table>
<thead>
<tr>
<th>Message</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficacy</td>
<td>46</td>
<td>53</td>
<td>52</td>
<td>49</td>
<td>50</td>
<td>43</td>
</tr>
<tr>
<td>Antihistamine</td>
<td>23</td>
<td>23</td>
<td>29</td>
<td>24</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>Rapid Onset of Action</td>
<td>16</td>
<td>17</td>
<td>17</td>
<td>20</td>
<td>17</td>
<td>21</td>
</tr>
<tr>
<td>Nonsedating</td>
<td>16</td>
<td>10</td>
<td>18</td>
<td>21</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>Safety</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Seasonal Use</td>
<td>12</td>
<td>11</td>
<td>6</td>
<td>12</td>
<td>9</td>
<td>7</td>
</tr>
</tbody>
</table>

Schering’s expert interpreted this study to indicate that one of the most consistent ZYRTEC sales messages is that ZYRTEC is an effective “nonsedating” antihistamine with a rapid onset of action. Schering’s expert also interpreted the data to indicate that Pfizer

53. See Schering, 189 F.3d at 222-23.
54. Id.
representatives were communicating a “nonsedating” message at a rate of 15-20 percent over the five months of the survey.

2. Summary of Pfizer’s Substantive Criticisms of Studies

Pfizer offered several significant methodological criticisms of the various surveys on which Schering relied. Some criticisms were specific to a particular survey; for example, the Bartolomeo survey used a leading question format, which repeatedly referred to “sedation,” e.g., “You might have mentioned this already, but I just want to make sure I have this right. During [the] Zyrtec detail, did the rep say anything about Zyrtec tablets and sedation?” and “Even at the risk of repeating what you have already told me, please be as complete and specific as possible about what was said about Zyrtec tablets and sedation?” Pfizer argued that such suggestive questions may have led the respondent to believe “sedation” was addressed, whether or not the representative discussed the topic, and telegraphed to respondents that their responses to the questions should be phrased in terms of “sedating/nonsedating.”

But the most significant methodological criticisms applied to all five of Schering’s surveys. First, none of them used any sort of control group which, Pfizer argued, prevented Schering from linking the survey response to the precise words that a Pfizer detail representative allegedly spoke. A physician who used the phrase “nonsedating” in a survey response might merely have reflected a conclusion that the physician inferred from the detail presentation and not something that Pfizer improperly stated or implied. Further, the presence of the phrase “nonsedating” in survey responses may reflect “noise,” i.e., a physician’s understanding based upon some influence not related to the detail presentation, such as medical journal articles, clinical experience with patients, or some other external stimuli or experience. Pfizer’s survey expert, Dr. Ross, testified that a control group was essential to making the critical causal connection between the survey response and Pfizer’s detail representative.55

In addition, Pfizer argued that the surveys could not demonstrate whether physicians used the phrase “nonsedating” in the same dichotomous manner that Schering urged during the litigation. Pfizer pointed to verbatim responses in the Schering survey data to show that some doctors were obviously using the term “nonsedating” to refer to products that caused some degree of “drowsiness” or “somnolence.” The subtlety of this semantic issue was at the heart of the litigation. Did Pfizer sales representatives actually utter the word “nonsedating,” or did they present truthful

and nonmisleading information (i.e., the incidence data) from which the physician concluded that ZYRTEC could be considered essentially "nonsedating"? And, if so, was the physician using "nonsedating" as Schering did? Pfizer argued that Schering's surveys should have been designed with more sensitivity and precision so as to ensure that the answer reflected a statement prohibited from use by Pfizer as opposed to a statement that was permitted, such as the FDA labeling data.\textsuperscript{56} Dr. Ross testified that the surveys relied on by Schering are not capable of showing what Pfizer representatives said or implied, emphasizing that the surveys are "too blunt a measuring instrument to pick up that very precise, but subtle issue."\textsuperscript{57}

Finally, Pfizer argued that the apparently consistent (~15-20 percent) rates across the five studies simply reflected a consistent level of "noise" and could not be properly interpreted as adding credibility to any of the surveys.\textsuperscript{58} Dr. Ivan Ross explained that the "consistent" level of "nonsedating" responses likely would constitute the "irreducible minimum" level of playback attributable to stimuli external to the sales presentation.\textsuperscript{59}

3. Summary of Pfizer's Legal Objections

During the preliminary injunction hearing, Pfizer moved to exclude Schering's surveys as inadmissible hearsay—an out of court statement offered to prove the truth of the matter asserted\textsuperscript{60}—arguing that physician responses to surveys are inadmissible to prove the precise content of the Pfizer sales representatives' statements to physicians.\textsuperscript{61} Pfizer contended that

\begin{itemize}

\item \textsuperscript{57} See Schering, Prelim. Inj. Tr. 435:1-12.

\item \textsuperscript{58} See Schering, Prelim. Inj. Tr. 396:13-397:10, 398:16-399:18 (testimony of Pfizer survey expert Ivan Ross).

\item \textsuperscript{59} See Schering, Prelim. Inj. Tr. 398:16-399:18; 421:17-422:8.

\item \textsuperscript{60} Hearsay is "a statement, other than one made by the declarant testifying at the trial or hearing, offered in evidence to prove the truth of the matter asserted." Fed. R. Evid. 801(c); see Blue Cross and Blue Shield of New Jersey, Inc. v. Philip Morris, Inc., 141 F. Supp. 2d 320, 323 (E.D.N.Y. 2001) (hearsay is generally excluded because it is untested by cross-examination, which exposes imperfections of perception, memory, narration, and opinion).

\item \textsuperscript{61} See Schering, Prelim. Inj. Tr. 219:22-220:5 (motion of Pfizer's counsel to exclude surveys as inadmissible hearsay); see also bench memorandum in support of defendants' joint motion to exclude plaintiff's survey evidence. Schering Corp. v. Pfizer Inc. and UCB Pharma, Inc., 98 Civ. 7000 (LMM) (S.D.N.Y. Oct. 29, 1998).
\end{itemize}
the admissibility of the surveys could not be salvaged under the residual exception to the hearsay rule because the surveys could not be considered reliable or trustworthy. Pfizer also argued that Schering was attempting to use surveys for a novel purpose, i.e., proving verbatim recall by physicians. In this situation, Pfizer urged that the surveys could not support expert testimony under Rule 703 of the Federal Rules of Evidence, which permits expert testimony based on inadmissible data only if such data routinely are relied upon by experts in the field for the same purpose.

Pfizer also argued that, even if the surveys and related expert testimony were admissible, the surveys’ methodological flaws, including most notably the lack of control and failure to clarify ambiguous responses, indicated that they should be entitled to limited or no weight and should not support the drastic remedy of a preliminary injunction. Pfizer argued that Schering had no likelihood of success on the merits because Schering’s evidence failed to establish that Pfizer’s advertising statement deceived a substantial portion of physicians who prescribe antihistamines.

4. Initial District Court Ruling

The district court excluded Schering’s survey evidence as hearsay that could not be admitted under any exception to the hearsay rule. Because Schering’s case rested almost exclusively on the survey evidence, the district court denied Schering’s motion for a preliminary injunction due to insufficient evidence. The court recognized that surveys are routinely accepted into evidence in Lanham Act cases to show respondents’ present sense perceptions and reactions to “advertising the content of which has been independently proved—by the introduction in evidence, say, of printed materials.” In those cases, the surveys reflect respondents’ “presently existing state of mind, attitude, or belief.”

62. See Yurman Design Inc. v. Diamonds and Time, 169 F. Supp. 2d 181, 185 (S.D.N.Y. 2001) (a preliminary injunction is “extraordinary,” “drastic,” and “should not be routinely granted” but only when the movant shows irreparable harm and either a likelihood of success on the merits or sufficiently serious questions on the merits to make them a fair ground for litigation, and a balance of hardships tipping decidedly in the movant’s favor).


65. Id. at *7; see also id. at *1 (holding that the record lacked “admissible evidence of what Pfizer’s and UCB’s representatives said to physicians [and therefore represents] an insufficient basis on which to find that Pfizer or UCB violated the settlement agreement as claimed”).

66. Id. at *2.
thus falling within the “state-of-mind” exception to the hearsay rule.\textsuperscript{67}

By contrast, Schering’s survey evidence was not offered to prove the physician-respondents’ mental impressions of a known advertisement or product; instead, they were introduced to prove “a fact remembered, i.e., what was said to them” by Pfizer’s sales representatives.\textsuperscript{68} The district court concluded that when surveys are introduced for this purpose, they constituted hearsay evidence, i.e., the physicians’ memories of the content of an advertising message (what was said by the sales representatives) which was in dispute, rather than the physicians’ mental impressions taken away from a proven advertising message.\textsuperscript{69} The court believed that allowing the use of survey data in this manner to establish the literal words spoken by the Pfizer sales representative would lead to a “virtual destruction of the hearsay rule.”\textsuperscript{70}

The court also considered Federal Rule of Evidence 703, which permits expert testimony based upon data that are not themselves admissible into evidence, if such data are “of a type reasonably relied upon by experts” in the field of surveys and market research.\textsuperscript{71} However, the court found that the use of surveys to demonstrate the content of a promotional message was entirely novel. Survey experts are not experts in “what one person said to another”\textsuperscript{72} and thus it could not be the case that a so-called expert would reasonably rely on surveys for that purpose. Finally, the court rejected Schering’s arguments that the expert testimony and survey data should come into evidence under the residual exception to the hearsay rule, Federal Rules of Evidence 807,\textsuperscript{73} and that a survey commissioned by Pfizer (the MMI off-the-shelf survey described above) was a party “admission.”\textsuperscript{74}

\begin{itemize}
  \item 67. Id. at *2-3, relying on Fed. R. Evid. 803(3), excepting from the hearsay rule a “statement of the declarant’s then existing state of mind, emotion, sensation, or physical condition (such as intent, plan, motive, design, mental feeling, pain, and bodily health), but not including a statement of memory or belief to prove the fact remembered or believed unless it relates to the execution, revocation, identification, or terms of declarant’s will.”
  \item 68. Id. at *3. The court found that the state-of-mind exception to the hearsay rule specifically excludes “statements of memory or belief to prove the fact remembered or believed.” See id. at *4.
  \item 69. Id. at *3.
  \item 70. Id. at *5 (quoting from Fed. R. Evid. 803(3) advisory committee’s note).
  \item 71. Fed. R. Evid. 703; see Schering, 1999 WL 144921, at *5.
  \item 72. Schering, 1999 WL 144921, at *5.
  \item 73. See id. at *3-4.
  \item 74. See Fed. R. Evid. 801(d)(2)(C). The court found that the survey commissioned by Pfizer was conducted by a survey firm independent of Pfizer and was not subject to Pfizer’s control. The court also found that an internal Pfizer memorandum referring to the survey was not admissible under this exception, as it merely summarized the report of the independent consultant. Schering, 1999 WL 144921, at *6.
\end{itemize}
5. The Court of Appeals Decision

On Schering’s expedited appeal, the Court of Appeals for the Second Circuit vacated the district court’s order, holding that the district court “relied on an erroneous per se rule against memory surveys offered to prove the facts remembered.” The appellate court remanded, instructing the district court “to determine the surveys’ trustworthiness on the basis of their methodological strengths and their relative susceptibilities to the risks of faulty memory and perception.” The Court of Appeals emphasized that, in addition to a claim of literal falsehood (which was all that the district court found Schering had asserted), Schering had now contended that the Pfizer sales representatives were “implying” that ZYRTEC was nonsedating. Therefore, those “main message” surveys were relevant to determine the “impressions” that the sales representative left with the physician. The appellate court explained that “plaintiffs alleging an implied falsehood are claiming that a statement, whatever its literal truth, has left an impression on the listener that conflicts with reality,” which “invites comparison of the impression, rather than the statement, with the truth.”

The court did not address two salient facts in reaching its conclusions: First, this was the first case in which surveys measured the impressions without knowing the precise content of the statement. Obviously, impressions alone are not actionable under the Lanham Act; it is only those impressions that are causally linked to a defendant’s utterance that give rise to a claim under the statute. Although implicit in the analysis, the Court of Appeals did not discuss this issue or the novelty of using survey evidence for this purpose. Second, the 1996 Settlement Agreement between the parties created safe harbors for sales representatives to convey indisputably truthful FDA labeling information to physicians (such as the fact that ZYRTEC’s incidence of sedation is between 11 to 14 percent in adults, and 2 to 4 percent in children). The court did not discuss the possibility that the FDA-approved data itself could cause some physicians to characterize ZYRTEC as “nonsedating.”

The court was careful to point out that its “ruling concerning admissibility in no way suggests that the district court should give these surveys any particular weight on remand.” On remand, the

75. Schering Corp. v. Pfizer Inc. and UCB Pharma, Inc., 189 F.3d 218, 221 (2d Cir. 1999).
76. Id.
77. Id. at 228.
78. Id. at 229 (emphasis added).
79. See Physicians’ Desk Reference, 2404-06, supra n.28.
80. Schering, 189 F.3d at 230.
district court might consider methodological criticisms raised by Pfizer, including the fact that the surveys were uncontrolled, the fact that Schering failed to call as witnesses any physicians to corroborate Schering’s allegations, and the fact that the FDA does not categorize antihistamines as either “sedating” or “nonsedating.” The Court of Appeals acknowledged that, under its holding, the surveys might be admitted into evidence, but receive no weight by virtue of their methodological flaws.

Turning to the “central reason” that Schering introduced its survey evidence, i.e., to show “literal falsehoods,” the Second Circuit agreed with the trial court that the state-of-mind exception to the hearsay rule81 plainly excludes any “statement of memory or belief to prove the fact[s] remembered or believed.”82 It concluded, however, that surveys offered to show literal falsehoods nevertheless could be admissible under the “residual hearsay rule.”83 Although the residual hearsay exception is to be used “rarely, and only in exceptional circumstances,”84 the Court of Appeals believed that the admissibility of the surveys could be salvaged under the exception, but only if the surveys could be shown to be trustworthy and methodologically sound, i.e., exhibiting “circumstantial guarantees of trustworthiness.”85 Once again faulting the lower court for “relying on an erroneous per se rule” (here, a rule against the use of out-of-court memory statements to “prove the facts remembered”),86 the appellate court indicated that the trial court should examine the surveys for “methodological strengths” and “proneness to faulty memory and perception” in order to assess admissibility under the residual hearsay exception.87

The Second Circuit vacated the use of the so-called “per se rules” against Schering’s memory surveys but expressly left it to the district court’s judgment to conclude “whether there are any other grounds for excluding” the surveys and “what weight to give all of the evidence admitted on remand.”88 Notwithstanding this

81. Fed. R. Evid. 803(3).
82. Id.; Schering, 189 F.3d at 227.
83. Fed. R. Evid. 807 (“A statement not specifically covered by Rule 803 or 804 but having equivalent circumstantial guarantees of trustworthiness, is not excluded by the hearsay rule, if the court determines that (A) the statement is offered as evidence of a material fact; (B) the statement is more probative on the point for which it is offered than any other evidence which the proponent can procure through reasonable efforts; and (C) the general purposes of these rules and the interests of justice will best be served by admission of the statement into evidence.”).
84. 189 F.3d at 232.
85. Id. at 238.
86. Id. at 221.
87. Id. at 234.
88. Id. at 240.
indication of deference to the district court, the Court of Appeals signaled that it believed the surveys should be admitted—and given weight—on remand. Without citing to the record, the Court of Appeals assumed that somnolence was a “critical factor” in physicians’ decision to prescribe ZYRTEC or CLARITIN and that the physicians surveyed “presumably would have been poised to look for and remember this type of information in a detailing.” It also assumed that “all five [of Schering’s] surveys tended to corroborate one another” based on the fact that all five surveys showed a similar magnitude (~15-20 percent) of doctors reporting that “nonsedating” was a message conveyed during the detail session. The Court of Appeals also indicated that the surveys were not unduly susceptible to faulty memory based upon “the fact that most of the surveys were conducted within a day, and one within a week, of [the detailing sessions].”

6. Post-Remand Decision of the District Court

On remand, the district court admitted the surveys into evidence, as suggested by the Second Circuit. It then gave them sufficient weight to grant, in part, Schering’s request for a preliminary injunction, preliminarily enjoining Pfizer from stating or implying that ZYRTEC is nonsedating or essentially nonsedating. The trial court denied Schering’s request for an order requiring Pfizer to disseminate corrective advertising in part because such advertising “could be interpreted to suggest, incorrectly, that the FDA has established a classification scheme regarding sedation of antihistamines, which is untrue.”

The district court’s reasoning was precise. Notwithstanding the holding by the Court of Appeals that Schering’s surveys were relevant to both literal and implied falsehood claims, on remand, the lower court evaluated the surveys only in connection with Schering’s “literal falsehood” case, holding that Schering’s surveys were not designed to elicit the facts necessary for an implied falsehood claim and that “any findings in this regard would be speculation.” The court held that Schering’s surveys were sufficiently trustworthy to support the preliminary injunction and

89. Id. at 236 (emphasis added).
90. Id.
91. Id. The Court of Appeals indicated that, on remand, one survey should be admitted as a party admission because a Pfizer employee approved the survey’s methodology before the research was performed and wrote a memo summarizing the results, which indicated to the court that the Pfizer employee “had to believe” that an improper message was being communicated by ZYRTEC sales representatives. Schering, 189 F.3d at 239.
93. Id. (internal quotations omitted).
94. Id.
rejected Pfizer’s arguments that the surveys were methodologically flawed and yielded ambiguous results. A control was not necessary, as it would have been in an implied falsehood case, because here, Schering was not trying to establish the “state of mind” of physicians but rather the “main message” conveyed by the Pfizer detail representatives.\(^95\) Hence, while a control would be necessary to make the “causal” connection between a literally truthful statement and an implied false belief, no control was needed to demonstrate a physician’s memory of a literally false promotional statement.

As for Pfizer’s argument that physicians may characterize the truthful FDA labeling data (e.g., that ZYRTEC has an incidence of somnolence between 11 and 14 percent) as meaning that ZYRTEC is “essentially nonsedating,” the court construed this possibility as an “unresponsive answer” to the survey question posed that was unlikely and, in any event, identifiable and thus excludable from the survey results.\(^96\) Finally, the court followed the lead of the Second Circuit in citing the “substantial consistency of the results of all five surveys.”\(^97\) The court concluded that all five of Schering’s surveys “possess[ed] sufficient weight to support a showing that a minority”\(^98\) “certainly less than 20%”\(^99\) “of the Zyrtec detailers, during the period March through August, 1998, were telling physicians, during detail visits, that Zyrtec was nonsedating or essentially nonsedating.”\(^100\)

### 7. Settlement and Permanent Injunction

About nine months after the district court’s decision on remand, the parties settled the litigation prior to trial. The preliminary injunction was converted to a permanent injunction that expressly provided that the parties were not bound by any interim judicial findings about survey methodology or trustworthiness.\(^101\)

---

95. Id. at *6.
96. Id. at *4.
97. Id. at *8.
98. Id.
99. Id.
100. Id.
101. Schering Corp. v. Pfizer Inc. and UCB Pharma, Inc., 98 Civ. 7000 (LMM) (S.D.N.Y. Apr. 20, 2001) (order granting permanent injunction). The injunction states: “Nothing contained in this Injunction shall limit any party’s right to challenge the accuracy, reliability, validity, or significance of any survey proffered by any other party in support of an alleged violation of this injunction.” Id. ¶ 4.
III. SUBSEQUENT SCIENTIFIC DATA DEMONSTRATES THAT SCHERING'S LITIGATION SURVEYS WERE UNRELIABLE

As part of its preparation for a trial on the merits, in 2000 Pfizer commissioned two surveys scientifically designed to test the central arguments that it had made during the preliminary injunction phase of the litigation. One survey sought to determine whether physicians understand the sedation side effect as a spectrum, in which ZYRTEC, although it produces somnolence, might nevertheless be characterized by some physicians as “nonsedating.” The other tested whether physicians would state, in response to survey questions, that a sales presentation indicated that a product was “nonsedating,” even though the sales presentation never actually used that term and only contains information expressly permitted by the 1996 Settlement Agreement between the parties.

The two surveys discussed below were designed and implemented with these hypotheses in mind. These surveys were complex and time-consuming both to design and to implement. For example, they involved interviewing almost 1,000 physicians who prescribe antihistamines as part of their practice; a large sample of “relevant” physicians was needed so that the data could be projectable to the total population of those physicians likely to be influenced by antihistamine promotion. Accordingly, this type of evidence would have been extremely difficult to develop during an expedited discovery schedule in advance of a preliminary injunction hearing. Because the litigation settled in advance of a final hearing on the permanent injunction, no court ever considered these data. The studies are presented here to demonstrate the problems that may arise when courts that lack scientific expertise are called upon to make decisions on a severely truncated evidentiary record. The surveys discussed below appeared in the Spring 2002 issue of Marketing Research.103

102. Bruce P. Keller et al., Surveys in False Advertising Cases, 624 PLI/Pat 351, 368-69 (2000) (“The demographics of [a] sample population must closely approximate those of the universe. If not, it will be difficult to persuade a judge that the responses of the sample can be projected to the relevant population at large. . . . Another aspect of projectability is the statistical reliability of the method used to select sample members. ‘Mall intercept’ surveys, in which consumers are randomly selected at various shopping malls, are, in theory, less statistically projectable than ‘probability surveys,’ in which each participant has a known chance of selection from the universe.”).

103. Wind, Marketing Research in the Courtroom, supra n.8.
A. Experiment No. 1: How Do Physicians Use the Term “Nonsedating”?  

In June through July 2000, Wharton marketing professor Dr. Yoram (Jerry) Wind designed and supervised a survey of 399 internists, general practitioners, family practitioners, and allergists who prescribe antihistamine products as part of their practice. The survey respondents were interviewed in person and presented with small white cards that displayed the brand name logos of five antihistamines: ALLEGRA, BENADRYL, CHLOR-TRIMETON, CLARITIN, and ZYRTEC. The physicians were asked to place the five brands, along with a card labeled “placebo,” on a visual analog scale (VAS) ranging from “Not Sedating At All” to “Extremely Sedating.” The VAS technique is an accepted method of evaluation that permits a researcher to make quantitative measurements along a continuum, particularly when the test variable can be measured by degrees of intensity or in relation to other variables. As used in this case, VAS presents physicians with a measuring device capable of recording their assessment of the relative incidence of somnolence among various antihistamine products.

Dr. Wind’s interviewers first asked the surveyed physicians the following: “On the scale please place the five products and a placebo in a way that best reflects your perception of their position on that scale.” After the five product and placebo cards were individually located on the VAS, the physicians were then asked to “draw a line that would separate sedating products from the nonsedating products.” No definitions of “sedating” or “nonsedating” were provided, so that each physician would utilize his/her own definition by where he/she drew the line. Finally, the interviewers asked respondents, “How would you describe a
nonsedating antihistamine?” The objective here was to capture whether physicians would articulate Schering’s “no different from placebo” dichotomous classification.

Physical marks/placements on the VAS allow a researcher to later assign numerical values to such placements. To do this, Dr. Wind’s team placed an overlay on the VAS responses that contained numbered, evenly divided hash marks from 1 to 10. The researchers then could assign a numerical value for each survey participant’s placement of each product as well as the physicians’ sedating/nonsedating demarcation lines. The values assigned for the five products, placebo, and the sedation line were averaged and a mean value assigned. The results were:

The results show a number of interesting findings:

1. The 399 physicians surveyed perceived that CLARITIN and ALLEGRA are no more (or less) sedating than a placebo. Stated differently, the survey shows no statistical difference between the mean values of these three items. This perception is consistent with the FDA-approved labeling for these products.

2. The physicians also understand that those antihistamines that are labeled as causing drowsiness or somnolence (ZYRTEC, BENADRYL, and CHLOR-TRIMETON) do so. All were ranked as more “sedating” than placebo, CLARITIN, and ALLEGRA by statistically significant margins. Moreover, the physicians perceive real differences among the three in the degree or extent of sedation caused. Each of the three is statistically different from each other.

3. The physicians draw the “average” line between “sedating” and “nonsedating” at a point approximately congruent with the average value for ZYRTEC.

Because averages may conceal important subgroups, Dr. Wind did a second analysis by separating three segments: (1) those physicians who perceived ZYRTEC to be among the “nonsedating” brands (i.e., those to the left of the location where the physician drew his/her sedating/nonsedating line); (2) those physicians who perceived ZYRTEC on the line separating “sedating” and
“nonsedating” products (i.e., did not clearly put ZYRTEC on either side of the line); and (3) those physicians who perceived ZYRTEC among the “sedating” brands (i.e., those to the right of the sedating/nonsedating line). The array of products within each segment is as follows:

This refinement produces more interesting findings:

1. Physicians did not differ widely in placing the nonsedating/sedating line, whether they viewed ZYRTEC on one side of it or the other. The differences among 4.69, 3.81 and 3.73 were not statistically significant.

2. Even when they differed about placing ZYRTEC above, on, or below the nonsedating/sedating dividing line, physicians were remarkably consistent in their perceptions of the placement of all products on the scale. In each segment, CLARITIN and ALLEGRA were viewed as equivalent to placebo, and the remaining products were in the same order, all statistically higher from placebo/CLARITIN/ALLEGRA and from each other.

3. Regarding the placement of ZYRTEC, the physicians disagreed, with roughly 47 percent describing it among the nonsedating drugs, 44 percent describing it among the sedating, and 9 percent drawing the dividing line at the point they put ZYRTEC. Notably, however, in each group—even those who described ZYRTEC as
nonsedating—the mean value for ZYRTEC on the scale was still statistically different from the values for placebo, CLARITIN, and ALLEGRA.

In other words, all of the physicians recognized that ZYRTEC was more likely to cause somnolence than placebo, including the substantial proportion (47 percent) of physicians who described ZYRTEC as “nonsedating.” This finding makes it improper to infer from a physician’s report that “Zyrtec is nonsedating” that the physician is misinformed, confused, or deceived about the differences in this side effect between ZYRTEC and placebo or CLARITIN. Physicians who juxtaposed ZYRTEC and “nonsedating” did not misunderstand the differences between ZYRTEC and placebo or CLARITIN.

The final question asked how the physician described a “nonsedating antihistamine.” Notably, none of the 399 respondents in the survey articulated the definition that Schering used in the litigation, i.e., that nonsedating antihistamines are those antihistamines that cause an incidence of sedation that is not statistically different from placebo. The survey respondents talked either about degrees of performance impairment or the patient’s subjective feeling of lethargy. This finding indicates that this “sophisticated” audience of trained physicians probably does not use key terminology as Schering contends.

Pfizer and UCB Pharma commissioned a separate survey of a substantially similar design that verified the results of Professor Wind’s VAS survey. This study was conducted by Walter McCullough of Monroe Mendelsohn Research, Inc. and was completed four months after the Wind study, in November 2000. The two studies (McCullough’s and Wind’s) were conducted independently; that is, neither surveyor had knowledge of the existence of the other study. McCullough surveyed 303 physicians who prescribe antihistamines as part of their practice. The sample included physicians from the fields of internal medicine, family practice, general practice, ear-nose-throat, and allergy. To ensure geographic diversity, physicians were chosen from 19 geographically disbursed locations throughout the continental United States. A visual analog scale was presented to physicians that was approximately eight inches in length and had a left endpoint indicating “no sedation” and a right endpoint indicating “a very high level of sedation.” No numbers were used in the scale, and there were no marks between the endpoints. Instead of the logos used by Professor Wind, Mr. McCullough asked physicians simply to mark the scale for five antihistamines and placebo. After collecting the data, Mr. McCullough overlaid a 0-100 point numerical grid to correlate numerical values with the physicians’ placement selections along the visual analog scale. Mr. McCullough’s findings were strikingly similar to the findings from
Dr. Wind’s independent study. Mr. McCullough calculated the average values as follows: placebo (14.7); ALLEGRA (17.7); CLARITIN (18.3); ZYRTEC (36.8); sedation line (41.0); CHLOR-TRIMETON (71.5); and BENADRYL (82.5). Like Dr. Wind, Mr. McCullough’s study shows that, on average, the physicians’ rankings have placebo, ALLEGRA, and CLARITIN clustered close together on the nonsedating side of the scale; ZYRTEC is to a measurable degree to the right of these, and very close to the “sedation line”; and CHLOR-TRIMETON and BENADRYL fall on the “sedation” side of the scale. Mr. McCullough further reported that 37 percent of the surveyed physicians ranked ZYRTEC on the “sedating” (left) side of the line, 62 percent ranked ZYRTEC on the “nonsedating” (right) side of the line, and 1 percent placed ZYRTEC at the line.

Although the Wind and McCullough studies provide valuable insight into physicians’ terminology and understanding of the relative incidence of somnolence among popular OTC and prescription antihistamines, it does not address a second critical issue, namely how doctors describe, in a survey setting, the contents of a detail presentation. To address this question, Pfizer commissioned a different study, discussed next.

B. Experiment No. 2: How Do Physicians Report the Content of a Drug Detailing Session?

During August and September 2000, Wharton marketing professor Paul E. Green designed and supervised a study of 578 general practitioners, family practitioners, internists, and allergists who prescribed antihistamines as part of their practice. The survey was implemented in the field by Guideline Research (New York, NY) under the direction of team leader Robert Reitter. Physicians were shown a videotape that depicted a sales representative having a detailing discussion with a doctor. About 24 to 48 hours after the physician saw the videotape, an interviewer contacted the physician and asked questions that were patterned after those used in Schering’s two surveys done for the litigation. The 24-48 hour delay between exposure to the stimulus and questioning was intended to mimic the survey methodology of the Schering surveys, which identified respondents by asking physicians if they had been detailed on ZYRTEC within the previous one to two days.

The videotape approach offered several advantages. First, it assured a standard oral and visual presentation, analogous to a print advertisement that could be read with no variations in tone,

109. Mr. McCullough’s report is on file with the authors.
110. Wind, Marketing Research in the Courtroom, supra n.8.
wording, demeanor, etc. that might occur with different live presentations. Second, the videotape could be presented in court, so that a judge could see exactly what the physicians saw.

Third, and most importantly, the videotape approach permitted the use of control groups. By making small changes in the script, Dr. Green could test the effects of these changes. For example, one concern was that physicians already had awareness of ZYRTEC, which had been marketed for over four years when the study was conducted. In addition, physicians also could be biased by awareness of litigation between Schering and Pfizer. To test whether prior awareness affected physician responses, some scripts substituted a hypothetical “new antihistamine” (called “Ardovac”), but the claims made were identical to those associated with ZYRTEC.

Another way in which controls were used was by varying the content of the script’s discussion of side effects. In this manner, Dr. Green could assess whether the extent of disclosures affected physician responses. Three levels of discussion were used. In each, the sales representative recited FDA-approved labeling information relating to somnolence for ZYRTEC. The variations were only in length (3-6 minutes) and amount of detail. The scripts contained only claims permitted under the 1996 Settlement Agreement between the parties, and the representative never used the words “nonsedating” or “sedating.” The same person delivered the sales presentation in each videotape. Each physician in the study saw only one videotape, which was selected at random according to a method to assure equivalent numbers of physicians in each cell.111

In sum, there were six cells, each with its own videotape, with the variations described above:

<table>
<thead>
<tr>
<th>(Amount of Detail in Video Presentation)</th>
<th>ZYRTEC</th>
<th>ARDOVAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Least</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Middle</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Most</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

One or two days after seeing the videotape, Dr. Green’s interviewers contacted the physicians by telephone. The physicians were asked: “What was said about (Zyrtec/Ardovac)?” “What, if anything, was said in the video about (Zyrtec/Ardovac) and side effects?”; and “What, if anything, was said in the video about (Zyrtec/Ardovac) and sedation?”112 The questions mimicked those asked in the Schering litigation survey. The “verbatim” answers

111. Id.
112. Id.
given by the physicians to the interviewers were then coded by Dr. Green and Mr. Reitter, tabulated, and subjected to statistical analysis.

Dr. Green reached the following conclusions:

(1) Approximately 25.6 percent of survey respondents reported that the sales representative in the videotape said that the product was “nonsedating” or words to that effect (e.g., causes “no sedation”).

(2) There was no statistically significant difference between the ZYRTEC cells and the Ar dovac cells in terms of reporting that the sales representative had claimed the product was “nonsedating.”

(3) There were no statistically significant differences in responses based on the various levels of discussion of side effects. In all six cells, a substantial proportion of physicians reported the sales representative on the tape had said the product was “nonsedating.” Thus, the extent of commentary within the 3-6 minute framework did not affect responses.

The data from this study are strikingly comparable to those presented by Schering. The 25.6 percent level of “nonsedating” responses was of a similar magnitude to the ~15-20 percent of “nonsedating” responses across the five surveys found by Schering’s expert, Mr. Bartolomeo. By the reasoning of the circuit and district courts, such apparent consistency should add confidence to the inference of what was said from what was reported. The error in that reasoning is proven by this study, however. None of the tapes used the term “nonsedating.” Therefore, the survey response obviously could not be due to a stimulus (i.e., the detail session) that contained a prohibited message. This study demonstrates that it is scientifically incorrect to infer “what was said” (stimulus) from “what was reported.”

In short, the Green study demonstrates the scientific invalidity of the “scientific reasoning” used by the Second Circuit and the district court on remand.

C. Discussion of the Experiments

Taken together, the two experiments emphasize the hazards of using uncontrolled market research surveys to support deceptive advertising claims. Schering’s surveys may have provided interesting data relevant for certain purposes, but these data were inadequate to sustain the legal burden required in deceptive advertising litigation. Because the 1996 Settlement Agreement prohibits Pfizer from making certain statements about ZYRTEC’s drowsiness side effect (“prohibited statements”) but recognizes Pfizer’s right to make other statements about ZYRTEC’s
drowsiness side effect ("permitted statements"), Schering’s burden was to prove that Pfizer made prohibited (rather than permitted) statements. As demonstrated by the Green study, the evidence submitted by Schering did not satisfy Schering’s burden.

Pfizer’s surveys demonstrate that the Schering surveys did not reliably discern whether the “nonsedating” survey responses were caused by permitted or prohibited messages. Because a physician could have concluded, based on permitted statements made by a Pfizer representative (as well as other sources of information) that ZYRTEC is “nonsedating” or “essentially nonsedating” (as those ambiguous terms may be understood or used by the physician), Schering was obligated to causally tie its survey data to the actual words Pfizer sales representatives spoke to physicians.

For example, a physician who was induced by Schering to believe that all antihistamines must be categorized either as “nonsedating” or “sedating” could reasonably conclude that ZYRTEC is “nonsedating” rather than “sedating” based on FDA-approved labeling, which reports that most patients who take it (86 percent of adults and 95 percent of children) do not experience any drowsiness. Similarly, a physician correctly informed that FDA considers the side effects associated with ZYRTEC (including drowsiness) to be “mild or moderate” could reasonably conclude that ZYRTEC is “nonsedating” or “essentially nonsedating,” rather than “sedating.” In either case, Pfizer could not be held responsible in a lawsuit alleging false advertising for such judgments made by trained physicians. The Settlement Agreement, the Lanham Act, and FDA labeling requirements focus on the accuracy of the information conveyed by Pfizer to physicians. None of these dictate the conclusions physicians reach about ZYRTEC based on the dissemination of truthful and nonmisleading information (such as FDA labeling).

As noted, the Wind, McCullough, and Green surveys discussed above were obviously complex and time-consuming to administer. The expedited schedule ordered by the district court—which was not unusual and required that all discovery be completed within a three-week interval from the denial of the TRO to the preliminary injunction hearing—did not allow sufficient time to conduct “counter-surveys” before the decision to grant or deny the preliminary injunction. Pfizer did, however, retain a survey expert, Dr. Ivan Ross, to testify about factors of unreliability, such as the lack of controls, in Schering’s proffered survey. Dr. Ross explained that a control was needed because Schering sought to prove a cause-and-effect proposition, that is, a causal connection between Pfizer’s promotional message and physicians “take away” memory or perception of that message. A control would have isolated the doctors’ memory or perception of the marketing message from the physicians’ perceptions based on other stimuli not related to the
promotional message. A control was particularly necessary here because the surveyed audience was highly trained and sophisticated. Physicians are familiar with ZYRTEC, a well-known antihistamine, and likely have been exposed to substantial research, statistics, and other data, including their own and their patients’ experience with the drug.\textsuperscript{113}

The scientific literature was particularly relevant. Some respected medical treatises refer to ZYRTEC’s active ingredient, cetirizine, as “nonsedating” even though the treatises indicate that cetirizine is associated with drowsiness.\textsuperscript{114} Scores of medical journal articles similarly refer to cetirizine as nonsedating\textsuperscript{115} or “low sedating,”\textsuperscript{116} or mention related concepts such as cetirizine having no impact on the central nervous system\textsuperscript{117} or no impairment on performance of motor skills or cognitive function.\textsuperscript{118} The regulatory context adds to the confusion, with ZYRTEC classified or marketed as nonsedating in Europe. As noted, the FDA stated that “any suggestion . . . that FDA has classified antihistamine drugs into categories of sedating or nonsedating would be false and/or misleading” because no such classification exists.\textsuperscript{119}

The district court also considered other warning signs about the trustworthiness of Schering’s survey conclusions, including those in the survey responses themselves. The surveys indicated

\begin{enumerate}
\item Dr. Ross explained that a survey designed to make causal conclusions must have a control to account for “noise.” Noise may be more likely in a survey about a well-known product such as ZYRTEC because physicians can be expected to integrate their own experiences and information from other sources into their responses.

\item See, e.g., The Royal Pharmaceutical Society, Martindale The Extra Pharmacopoeia 427 (1996) (listing cetirizine as a “nonsedating antihistamine” and recognizing that it causes some drowsiness).


\item See, e.g., Robert P. Harvey et al., Model for Outcomes Assessment of Antihistamine Use for Seasonal Allergic Rhinitis, 97 J. Allergy & Clinical Immunology 1233, 1234 (June 1996); F. Bonifazi et al., Comparative Study of Terfenadine and Cetirizine in Hay Fever: Assessment of Efficacy and Central Nervous System Affects, 5 J. Invest. Allergol. Clinical Immunology 40 (Jan./Feb. 1995).

\item Sept. 24, 1996 letter to Ronald J. Garutti, Schering Corp., from Joan Hankin, Reg. Review Officer, FDA.
\end{enumerate}
that physicians responded based on their own experience with ZYRTEC or other extraneous information;\textsuperscript{120} that physicians were reporting their conclusions, which did not necessarily reflect the terminology of the representative;\textsuperscript{121} and, most importantly, that physicians did not use the term “nonsedating” as Schering did during the litigation. On this last point, the surveys indicated that physicians often use the term “nonsedating” when describing a drug that causes some amount of drowsiness.\textsuperscript{122} Thus, the survey responses demonstrated that when a physician says “nonsedating,” the physician does not necessarily mean that no drowsiness is associated with the drug. This ambiguity goes to the very heart of the issue in the litigation: if doctors use the term “nonsedating” to include some “somnolence” or “drowsiness,” then doctors’ “nonsedating” response cannot be used as evidence that Pfizer made a prohibited statement.

The surveys also indicated that physicians had difficulty recalling the specifics of what transpired and who said what to whom during the brief detail visits.\textsuperscript{123} Finally, Schering’s survey evidence was novel in that it measured reactions to verbal

\textsuperscript{120} The responses themselves demonstrate that physicians respond based on their own experience with ZYRTEC or other extraneous information. See, e.g., DX 33 (Dr. ID No. 10001) (“I know all about Zyrtec”); (Dr. ID No. 10006) (doctor told representative that tiredness occurred in only a small percentage of his patients); (Dr. ID No. 20063) (“I lecture on antihistamines and have full knowledge of Zyrtec.”); (Dr. ID No. 20015) (“very few of my people have had sedation”); (Dr. ID No. 20028) (“we are familiar with the drug”); (Dr. ID No. 20064) (“sedation is an infrequent problem for me”); (Dr. ID No. 20070) (“I told him Zyrtec was great. I didn’t give him a chance to say much.”); (Dr. ID No. 20083) (Pfizer representative said “[t]here is a slightly higher sedation but I said to him that I personally didn’t think so”).

\textsuperscript{121} “She said it was the fastest acting nonsedating second generation antihistamine available.” However, in response to a follow-up question the same doctor clarifies that the Pfizer representative had made no such statement: “Really didn’t mention or address the sedation aspect of it. I know about it, but she didn’t discuss it.” DX 33 (emphasis added).

\textsuperscript{122} For example, one physician said that ZYRTEC “causes some drowsiness,” but “doesn’t cause sedation.” DX 33 (Dr. ID No. 10010). Another physician described ZYRTEC as: “Nonsedating. 10% sleepiness.” In another survey a doctor stated “Zyrtec is the fastest acting non-sedating antihistamine which also gives the longest relief in a one-a-day dose form. It only has 5% more sedation and is indicated in children age 6 and above.” PX 5, at S 467 (Dr. ID No. 3B00033) (emphasis added). This fact demonstrates that the inclusion of the term “nonsedating” in a survey response is not a surrogate for a violation of the Settlement Agreement. See, e.g., DX 25, at 17-001005 (“For some patients that I have prescribed it to, it happens to be sedating, for others it is nonsedating”).

\textsuperscript{123} See DX 33 (Dr. ID No. 10002) (“see so many reps”); (Dr. ID No. 10008) (“I don’t remember that much about it. I wasn’t paying much attention.”); (Dr. ID No. 10014) (“I can’t remember for sure”); (Dr. ID No. 20026) (“Can’t recall”); (Dr. ID No. 20032) (“I was busy”); (Dr. ID No. 20036) (“I was in a hurry because of time constraints”); (Dr. ID No. 20085) (“I don’t remember what they all say”); (Dr. ID No. 20112) (“had to busy a day”); (Dr. ID No. 20128) (“don’t recall”; “I see so many reps” “but I am aware that Zyrtec tablets have minimum sedation”); (Dr. ID No. 20134) (“I can’t remember anything”); (Dr. ID No. 20150) (“Oh, I’m not paying so much attention to him”); (Dr. ID No. 20170) (“I’m really not sure what he said. I had a string of reps in yesterday—8-10 reps”).
statements. In other words, Schering was using survey data to identify the “stimulus,” that is, the content of the verbal statement itself, and the deceptive potential of the stimulus at the same time. This unprecedented application of survey methods suggested an additional reason to be concerned about the reliability of the conclusions presented.

The warning signs in Schering v. Pfizer indicated substantial relevance and reliability concerns about Schering’s survey evidence. The district court—which had the benefit of hearing testimony from the parties’ survey experts—initially heeded these warning signs and excluded the surveys from evidence. The Second Circuit’s opinion, vacating the district court’s decision, failed to analyze these factors. Instead, the Second Circuit appeared to assume that the “smoke” created by the surveys signaled the presence of fire, i.e., deceptive advertising. The Wind, McCullough, and Green surveys provide a scientific basis to conclude that the district court had it right the first time.

IV. BEYOND SCHERING v. PFIZER: PLAINTIFFS’ SCIENTIFIC BURDENS IN DECEPTIVE ADVERTISING CASES

A. Judicial “Gatekeeping” and Related Obligations to Insist on Sound Science

As with all expert evidence, proffered survey evidence in deceptive advertising cases must submit to the district court’s “gatekeeping” process. Gatekeeping is an obligation imposed by Daubert v. Merrell Dow Pharmaceuticals and its progeny to assess and exclude expert evidence that is not both “relevant and reliable.” This means that neither surveys introduced to show deceptive advertising nor related expert testimony may be admitted into evidence unless they can be shown to be derived

---

124. Schering, Prelim. Inj. Tr. (Bartolomeo) 173:9-174:21 (Mr. Bartolomeo admitted he had “never before used a survey . . . of a physician in order to determine and testify as to what a third party had told that physician who in turn told . . . [the] telephone interviewers”); Wetzel Dep. 74:15-21 (DTW survey chief Wetzel testified that he had “no way of knowing what the sales representative said to the doctor”); Wetzel Dep. 77:2-6 (DTW surveys “did not measure what the sales representative said to the doctor”).


126. Id. at 589.

127. See, e.g., SMS Sys. Maintenance Servs., Inc. v. Digital Equip. Corp., 188 F.3d 11, 25 (1st Cir. 1999) (“Expert opinions . . . are no better than the data and methodology that undergird them. . . . [A]n expert must vouchsafe the reliability of the data on which he relies and explain how the cumulation of that data was consistent with standards of the expert’s profession.”); see also Simon Property Group L.P. v. mySimon, Inc., 104 F. Supp. 2d 1033, 1039 (S.D. Ind. 2000) (“Consumer survey results must be presented through expert witnesses. Under Daubert, the gatekeeping function requires the court to ensure that expert testimony ‘is not only relevant, but reliable.’”) (citations omitted).
from the “scientific method”\textsuperscript{128} and to be reliably applied to the facts at issue.\textsuperscript{129} The Supreme Court in Daubert set forth a nonexclusive list of factors that can aid the analysis, such as whether a proposed expert study has a known or measurable rate of error;\textsuperscript{130} however, the Court emphasized that the gatekeeping inquiry is highly contextual and requires courts to ensure that “an expert . . . employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.”\textsuperscript{131} The list of relevant disciplines applicable to Lanham Act surveys obviously would include substantive areas such as marketing and consumer behavior; however, the list should also include areas such as statistics and scientific experimentation.\textsuperscript{132} Standards by which to judge the soundness of survey design have evolved over the past decade and a body of scientific and legal precedent now is available to aid a court’s gatekeeping function in deception cases.\textsuperscript{133}

\textsuperscript{128} See Daubert, 509 U.S. at 590.

\textsuperscript{129} Id. at 591-94. In performing the gatekeeping function, courts “must recognize that due to the difficulty of evaluating their testimony, expert witnesses have the potential to ‘be both powerful and quite misleading.’” Cooper v. Smith & Nephew, Inc., 259 F.3d 194, 199 (4th Cir. 2001) (citations omitted).

\textsuperscript{130} Id., 509 U.S. at 593-94 (factors to consider in determining whether proffered expert evidence is relevant and reliable can include (1) whether the expert’s theory or technique has been tested; (2) whether the expert’s theory or technique has been subjected to peer review; (3) the known or potential rate of error; (4) the existence and maintaining of standards controlling the technique’s operation; and (5) general acceptance in the scientific community). These factors are not always useful; cf. Peter Bronsteen and Asim Varma, Daubert Rules for Economists, 15 Antitrust, Summer 2001, at 14 (commenting that the Daubert factors offer “cold comfort in antitrust cases” because “some sound economic analysis . . . may not be empirically testable or have a known or measurable error . . . may not be of sufficient interest to merit publication or peer review . . . [or] may be too esoteric to have attracted widespread acceptance”).


\textsuperscript{132} James E. Clevenger, False Advertising Under Lanham Act, 44 Am. Jur. Proof of Facts 3d 1 (1997) (indicating that courts should assess surveys to ensure, inter alia, that “the sample, the questionnaire, and the interviewing were in accordance with generally accepted standards of objective procedure and statistics in the field of such surveys”); see also Shari Seidman Diamond, Reference Guide on Survey Research, in Reference Manual on Scientific Evidence 229, 238 (Federal Judicial Center ed., 2d ed. 2000) [hereinafter “Diamond, Reference Guide on Survey Research”] (“Experts prepared to design, conduct, and analyze a survey generally should have graduate training in psychology (especially social, cognitive, or consumer psychology), sociology, marketing, communication sciences, statistics, or a related discipline; that training should include courses in survey research methods, sampling, measurement, interviewing, and statistics.”).

\textsuperscript{133} The factors that guide a federal court’s evaluation of surveys in Lanham Act cases have become increasingly sophisticated. Numerous decisions have addressed various factors to consider in evaluating survey evidence. A court’s analysis of the relevance and reliability to be afforded proffered surveys may include consideration of the following questions: (1) whether the surveys can address the question at issue in the litigation; (2) whether the experts who designed, conducted, and/or analyzed the surveys are appropriately skilled, experienced, and unbiased; (3) whether the survey respondents were representative and projectable to the relevant universe; (4) whether the survey questions were clear, precise, and nonleading; (5) whether the surveys included an appropriate control group or control
The court’s gatekeeping function requires a close “fit” between the underlying data and the expert conclusion. A court has the authority to “conclude that there is simply too great an analytical gap between the data and the opinion proffered.” Whether there is a “fit,” the Court explained in Daubert, is not always obvious, and “scientific validity for one purpose is not necessarily scientific validity for other, unrelated purposes.” For example, offering a survey to show the existence of an opinion among relevant consumers is quite different from offering a survey to show what caused that opinion. A survey may be reliable for one purpose, but not the other.

The gatekeeping obligation often can be discharged during a single pretrial challenge to the admissibility of proffered expert testimony. Because deceptive advertising cases typically involve a request for interim injunctive relief (to be determined quickly) and permanent injunctive relief and damages (to be determined at trial), the Daubert inquiry may be repeated at each stage of the litigation. The gatekeeping function is not necessarily complete merely because a federal court provisionally admits or excludes survey testimony for purposes of a preliminary injunction proceeding; to the contrary, courts continue to evaluate the reliability and relevance of proffered expert evidence, among other things, to make an admissibility determination for purposes of a final hearing on the merits.

question; (6) whether the surveys were designed to neutralize bias; (7) whether the surveys were designed to neutralize faulty memory; (8) whether the survey responses were ambiguous and, if so, whether further, clarifying questions were used; (9) whether the results of the survey were fairly analyzed, tabulated, and reported; and (10) whether the application of the survey results to the issues in this case were fair and reasonable. See Phyllis J. Welter, Trademark Surveys (1997); Diamond, Reference Guide on Survey Research, supra n.132, at 233-35; Jacob Jacoby et al., Survey Evidence in Deceptive Advertising Cases Under the Lanham Act: An Historical Review of Comments from the Bench, 84 TMR 541 (1994); see also Alex Simonson, Survey Design in False Advertising Cases, 1207 PLI/Corp 309 (2000); Bruce P. Keller, et al., Surveys in False Advertising Cases, 624 PLI/Pat 351 (2000).

135. Daubert, 509 U.S. at 591.
136. See, e.g., Lanvin, Inc. v. Colonia, Inc., 776 F. Supp. 125, 127 (S.D.N.Y. Oct. 28, 1991) (holding that preliminary injunction “findings are not binding, particularly where, as here, hearsay evidence was introduced at the preliminary injunction hearing but would not be admissible in the summary judgment context”). In general, determinations made at a preliminary injunction stage are not binding on later proceedings, in part because such determinations “are often based on incomplete evidence and a hurried consideration of the
In Kumho Tire v. Carmichael,138 the Supreme Court confirmed that the “gatekeeping” obligation applies not only to testimony based on “scientific” knowledge “but also to testimony based on ‘technical’ and ‘other specialized’ knowledge.”139 Thus, to whatever extent there was doubt about whether Daubert—traditionally described as the standard for evaluating “scientific” expert evidence—applied to Lanham Act surveys, Kumho Tire clearly removed any such doubts.140 Courts must subject surveys to a Daubert analysis. Moreover, this legal rule is justified by science—many of the same principles from the “scientific method” used in the laboratories of physical and medical science are fully applicable to assessing the relevance and reliability of proffered Lanham Act surveys.

In another development since the Second Circuit decided Schering, the Daubert rule has been codified in an amendment to Federal Rule of Evidence 702, which became effective on December 1, 2000. The rule requires expert testimony to be (1) “based upon sufficient facts or data”; (2) to be “the product of reliable principles and methods”; and (3) the principles and methods must be “applied . . . reliably to the facts of the case.”141

The foregoing description of the Daubert requirements assumes that the proffered expert evidence is otherwise admissible. For example, the expert evidence is inadmissible if the proffered expert is not qualified to express the opinion offered,142 if the conclusion is based more on personal belief, speculation, or

---


140. Recently Lanham Act courts have relied expressly on Daubert and Kumho Tire in evaluating surveys. See Simon Property Group L.P. v. mySimon, Inc., 104 F. Supp. 2d 1033 (S.D. Ind. June 7, 2000); National Football League Properties, Inc. v. ProStyle, Inc., 57 F. Supp. 2d 665 (E.D. Wis. 1999); but see Bacardi & Co. Ltd. v. New York Lighter Co., Inc., No. 97-CV-7140, 2000 WL 298915, at *5 (E.D.N.Y. Mar. 15, 2000) (citing Kumho Tire, but questioning its applicability and instead relying on the arguably contradictory rule in Schering that “methodological attacks . . . go to the weight of such evidence, not to its admissibility.”); see also supra n.5. The Second Circuit in Schering did not mention the “gatekeeping” function, the Daubert rule, or the Supreme Court’s decision in Kumho Tire, decided five months earlier.


142. A court must look beyond the general qualifications of a proposed witness to determine whether he or she possesses the special knowledge or experience necessary to express a specific opinion to be presented in the context of the facts of the case. See Kumho Tire, 526 U.S. at 155-58.
impression rather than specialized knowledge, or if its probative value is substantially outweighed by the danger of unfair confusion or prejudice. Indeed, on this last point, a few months before the Second Circuit decision in Schering, a different panel of that Circuit rejected the appellant’s argument that “soundness of its survey method should go to the weight of this evidence, not its admissibility.” In that case, the appeals court affirmed the exclusion of a survey in a Lanham Act case on the basis that it was a “memory test” and, as such, “any probative value of the survey was outweighed by its potential to confuse the issues in the case.”

In deceptive advertising cases, federal courts often test the relevance and reliability of scientific evidence with regard to traditional scientific disciplines such as chemistry, medicine, and physics. For example, in so-called “literal” falsehood cases, which do not typically involve consumer surveys, litigants frequently dispute the truth or falsity of the advertising. This battle is often fought with “hard” science. Introduction of scientific evidence in these settings is used to prove the truth or falsity of the advertising statement. The challenged claim must be measurable as true or false. Plaintiffs must demonstrate a false or misleading fact capable of measurement; generalized assertions, opinions, or subjective claims that cannot be proven true or false—for example, “You’re in good hands with Allstate” or Papa John’s “Better Ingredients, Better Pizza”—are “puffing,” not actionable under the Lanham Act. Stated differently, the plaintiff must show that the alleged deception concerns a fact that is “material” to

---

143. Daubert, 509 U.S. at 589-90 (holding that, under Fed. R. Evid. 702, an expert may only opine as to scientific, technical, or other specialized knowledge and not “subjective belief or unsupported speculation”); see also Weisgram v. Marley Co. 169 F.3d 514, 519 (8th Cir. 1999), aff’d, 528 U.S. 440 (2000); Target Mkt. Publ’g, Inc. v. Adv’n, Inc., 136 F.3d 1139, 1143 (7th Cir. 1998); Greensboro Prof’l Firefighters Ass’n v. City of Greensboro, 64 F.3d 962, 967 (4th Cir. 1995); Lithuanian Commerce Corp., Ltd. v. Sara Lee Hosey, 179 F.R.D. 450 (D.N.J. 1998) (granting motion in limine excluding opinions of expert based on speculative assertions).

144. See Fed. R. Evid. 403; see Arche, Inc. v. Azaleia, Inc., 882 F. Supp. 334, 336 (S.D.N.Y. 1995) (“While methodological defects in surveys usually go to the weight rather than the admissibility of the evidence, there comes a point where the probative value of the survey is exceeded substantially by its prejudicial effect and potential for confusion and waste of time.”) (excluding purported survey evidence).


146. Id.


149. Pizza Hut, Inc. v. Papa John’s Int’l, Inc., 227 F.3d 489, 498-99 (5th Cir. 2000) (finding slogan to be puffery when alone, but actionable in combination with advertisements claiming that ingredients were fresher and better than those of Pizza Hut).

150. Pizza Hut, 227 F.3d at 504.
consumers’ purchasing decisions.\textsuperscript{151} When the methodology for measuring this relevant factual statement is disputed, the court takes up its “gatekeeping” role. It thus must be willing to delve into and decipher technical data and testimony and resolve the dispute with reference to the applicable scientific disciplines.

Many literal falsehood cases involve allegations that the defendant has misrepresented its product’s composition or efficacy. For example, in Transclean Corporation v. Bridgewood Services, Inc.,\textsuperscript{152} the plaintiff’s expert chemist relied on a “nickel tracer analysis” to disprove the defendant’s advertising claim that its automatic transmission fluid replacement system replaced “100 percent,” “all,” and “every drop” of the transmission fluid in a vehicle’s engine.\textsuperscript{153} The nickel tracer analysis survived a Daubert challenge at the summary judgment stage of the case—even though the test was created for purposes of the litigation. The district court noted that “[t]he Daubert factors have an unusual application in the milieu of testing advertising claims such as these because, often, the only need for these tests arises in litigation, and the tests are thus extemporized, to some extent, from the standardized methodology.”\textsuperscript{154} Nevertheless, the court analyzed the proffered expert evidence, determined that the experiment was patterned on reliable methodologies in the relevant industry, and concluded that the chemist likely would be able to discharge his obligation under Daubert to “employ[] in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.”\textsuperscript{155} The court warned that its determination was preliminary and that the defendant was free to raise again the Daubert challenge as a motion “in limine, and, in that event, we will revisit the issue as the facts and law require.”\textsuperscript{156}

Issues regarding the reliability and relevance of scientific testing also arise in the context of “tests prove”\textsuperscript{157} advertising claims. If an advertiser attributes a statement about superiority, efficacy, or product attributes to a “test,” a plaintiff may enjoin such claims under the Lanham Act by proving such “tests . . . were not sufficiently reliable to permit one to conclude with reasonable

\textsuperscript{151} Southland Sod Farms v. Stover Seed Co., 108 F.3d 1134, 1145 (9th Cir. 1997) (“A specific and measurable advertisement claim of product superiority based on product testing is not puffery.”).

\textsuperscript{152} 77 F. Supp. 2d 1045 (D. Minn. 1999).

\textsuperscript{153} Id. at 1098.

\textsuperscript{154} Id. at 1099.

\textsuperscript{155} Id. (quoting Kumho Tire Co. v. Carmichael, 526 U.S. 137, 152 (1999)).

\textsuperscript{156} Id.

\textsuperscript{157} See, e.g., United Indus. Corp. v. Clorox Co., 140 F.3d 1175, 1181-82 (8th Cir. 1998); BASF Corp. v. Old World Trading Co., 41 F.3d 1081, 1090-91 (7th Cir. 1994); Castrol, Inc. v. Quaker State Corp., 977 F.2d 57, 63 (2d Cir. 1992).
certainty that they established the proposition for which they were cited. Under this rule, Castrol successfully challenged Pennzoil’s superiority claim that its motor oil “outperforms any leading motor oil against viscosity breakdown.” Pennzoil based this claim on research Pennzoil conducted utilizing a testing procedure promulgated by the American Society of Testing and Materials (ASTM). The court was required to understand and apply the science of polymer viscosity in order to determine the limitations of the ASTM test:

This [ASTM] test was never intended to compare the viscosity breakdown of oils of different polymer classes, and the test cannot perform this function accurately. Pennzoil and Castrol are motor oils of different polymer classes, and thus this test’s comparison of the two oils proves nothing relevant. Actually, the test does not measure viscosity breakdown at all; rather, it measures percentage of viscosity loss. Thus, while Pennzoil’s ASTM test may have been reliable for some purposes, it was not a valid measure of viscosity breakdown and therefore could not support Pennzoil’s superiority claim. The Third Circuit affirmed the district court’s holding that Pennzoil’s advertising claim was literally false and the entry of a permanent injunction.

In another deceptive advertising suit against a producer of “turfgrass seed” that claimed “50 percent less mowing,” the Ninth Circuit denied a Daubert motion to exclude plaintiff’s expert in turfgrass breeding, whose testimony was proffered to attack the reliability of the litigation tests performed by defendant. The plaintiff fended off the Daubert challenge by establishing that the test was conducted according to the “scientific method.” The court held that a “proponent of scientific evidence may satisfy its burden of establishing that the evidence is scientifically valid” by any of the following:

- showing that the evidence grew out of pre-litigation research,
- showing that the research upon which the evidence is based has been subjected to normal scientific scrutiny through peer review and publication, or explaining precisely how the conclusions were reached and pointing to some objective source to show that the conclusions are based on “scientific method,

160. Id.
161. 987 F.2d at 944.
162. Id.
163. Southland Sod Farms v. Stover Seed Co., 108 F.3d 1134, 1141-42 (9th Cir. 1997).
as it is practiced by (at least) a recognized minority of scientists in the[ ] field.”

These deceptive advertising cases illustrate the method by which federal courts evaluate scientific evidence. As discussed below, the approach to considering deceptive advertising surveys really is no different. Daubert’s techniques for assessing expert evidence based upon “hard” science are transferable—both as a matter of science and, certainly after the 1999 decision in Kumho Tire, as a matter of law—to evaluating survey evidence.

B. Surveys as Controlled Experiments

The issue of whether controls are required for surveys in deceptive advertising litigation is critical to the federal courts’ obligation to ensure that survey evidence is relevant and reliable. Inexplicably, this issue has failed to merit careful discussion in many cases. Part of the problem may be that courts do not always take the time to consider the purpose for which surveys are offered in Lanham Act cases. Unlike a “public opinion poll” which polls individuals about their opinions, attitudes, and beliefs, a Lanham Act survey is intended to show how the content of an advertisement affects or influences perceptions. This is a critical difference. Nevertheless, courts quite frequently equate “surveys” with “opinion polls.” Sometimes courts, including the Second Circuit in Schering v. Pfizer, discuss the admissibility of surveys in Lanham Act cases with reference to the Advisory Committee note to Federal Rule of Evidence 703 concerning “public opinion

164. Id. at 1141 (emphasis added).

165. See Diamond, Reference Guide on Survey Research, supra n.132, at 256 (“Most surveys...are not conducted to describe consumer beliefs. Instead, they are intended to show how a trademark or the content of a commercial influences respondents’ perceptions.”).


167. 189 F.3d at 226.
poll evidence” as a basis for expert testimony.\textsuperscript{168} Few courts describe deceptive advertising surveys as “experiments.” The term “experiment” is used principally in connection with the physical and biological sciences, whereas the term “survey” is often associated with public opinion measurement, for example, polls to predict the outcome of elections to public office or to gauge public sentiment about a current public policy issue. The implications are not merely semantic.

Opinion polls are directed towards the measurement of the extent and degree to which a relevant sub-population holds a belief or opinion at a given moment in time. The poll is not concerned with determining the cause of the opinion but simply intends to measure its existence and magnitude. Whether called a “survey” or “opinion poll,” its relevance and reliability must be judged in light of the purpose for which it is offered and within the unique facts and circumstances of the litigation. The true “opinion poll” (measuring a belief without regard to its cause) obviously can have relevant and reliable probative value in certain litigation contexts. For example, in a recent Seventh Circuit case, the court permitted a tobacco manufacturer to introduce a public opinion poll taken in the early 1950s to demonstrate public understanding about the health attributes and addictive qualities of cigarette smoking during that decade.\textsuperscript{169} In obscenity litigation, courts have considered opinion polls to be relevant and reliable evidence of community standards of decency.\textsuperscript{170} Opinion polls can be probative in criminal cases on the issue of whether a defendant can obtain a fair trial within a particular jurisdiction; thus, for example, in a recent New York case, a “public opinion survey” showed that a substantial portion of the potential jury pool had already made a determination of the defendant’s guilt, which led the court to grant defendant’s motion for a change of venue.\textsuperscript{171}

In deceptive advertising litigation, a true “opinion poll” may measure a false belief or perception but cannot establish a causal link between that belief and an advertising statement. The legal element of “causation” in deceptive advertising cases often is discussed in the context of damages after a statutory violation is

\textsuperscript{168} The Advisory Committee note to Rule 703 of the Federal Rules of Evidence, which provides that expert testimony may be based upon data “reasonably relied upon by experts in the particular field,” states that Fed. R. Evid. 703 offers a “satisfactory basis for ruling upon the admissibility of public opinion poll evidence. Attention is directed to the validity of the techniques employed rather than to relatively fruitless inquiries whether hearsay is involved.” Fed. R. Evid. 703 Advisory Committee note (emphasis added).

\textsuperscript{169} Insolia v. Philip Morris, Inc., 216 F.3d 596, 601-02 (7th Cir. 2000).


\textsuperscript{171} People v. Boss, 701 N.Y.S.2d 342 (App. Div. 1999); see also Keith v. Volpe, 858 F.2d 467, 479-81 (9th Cir. 1988) (survey to show statistics concerning respondents’ race, income, and housing preferences).
demonstrated. In order to obtain money damages, for example, a plaintiff typically is obligated to provide additional proof of causality between the deceptive advertising claim and diversion of sales, lost profits, or other indicia of money damages.\(^{172}\) This concept of “causation” is well recognized in deceptive advertising cases but differs from the concept of “causality” that relates to the underlying claim of liability and that is the principal subject of this article. Increasingly, courts in Lanham Act cases understand and have expressly required surveys to demonstrate a legal requirement of causation related to the liability phase. Acknowledging the requirement of causation in deceptive advertising litigation follows the fundamental tenet of law that a defendant should not be exposed to liability absent proof of a causal relationship between his conduct and an alleged wrong.\(^{173}\)

The foregoing discussion does not mean that a “poll” can have no probative value in deception litigation; just that, standing alone, it cannot function as dispositive proof of liability. For example, in Pfizer Inc. v. Miles,\(^{174}\) the court considered a survey in which “pharmacists were telephoned and asked about their perceptions regarding the similarities between Adalat CC and Procardia XL.”\(^{175}\) While such a survey undoubtedly provides potentially usable information, including information about physicians’ generalized perceptions, usage of terminology, understanding of the products and their differences, and the existence of misunderstandings and misinformation, the survey is insufficient to establish the causation element of a deception claim.

\(^{172}\) See, e.g., Hot Wax, Inc. v. Turtle Wax, Inc., 191 F.3d 813, 819-20 (7th Cir. 1999) (holding that “to recover money damages under the [Lanham] Act, a plaintiff must prove both actual damages and a causal link between defendant’s violation and those damages”); see generally Balance Dynamics Corp. v. Schmitt Indus., Inc., 204 F.3d 683, 692 (6th Cir. 2000) (holding that plaintiff could recover “damage control” costs for counter advertising without showing marketplace damages but could not recover damages for loss to goodwill or disgorgement of competitor’s profits without further evidence of a causal link between the challenged statements and damages sought). Even in cases limited to obtaining injunctive relief, where the products are not obviously in competition or where the defendant’s advertisements make no direct reference to a competitor’s products, a plaintiff must show some indication of a causal connection between the challenged advertising and harm to the plaintiff. See Telebrands Corp. v. Media Group, Inc., No. 97 Civ. 6768, 1997 WL 790576, at *3 (S.D.N.Y. Dec. 24, 1997) (holding that “[t]o demonstrate that defendant’s false advertising will cause irreparable harm, plaintiff must ‘offer something more than a mere subjective belief that [it] is likely to be injured as a result of the false advertising; [it] must submit proof which provides a reasonable basis for that belief’”) (quoting Coca-Cola Co. v. Tropicana Prods., Inc., 690 F.2d 312, 316 (2d Cir. 1982)).

\(^{173}\) In tort cases, the legal elements of cause in fact and proximate cause are well established. Under the traditional rule, statistical correlation alone is insufficient. “[P]roof that can provide direct and actual knowledge of the causal relationship between the defendant’s tortious conduct and the plaintiff’s injury” is required. In re Agent Orange Prod. Liab. Litig., 597 F. Supp. 740, 835 (E.D.N.Y. 1984) (quotations omitted), aff’d, 818 F.2d 145 (2d Cir. 1987).


\(^{175}\) Id. at 447.
As the court in Pfizer Inc. v. Miles explained, “the respondents were not shown any of Miles’ promotional materials, which makes it impossible to determine whether Miles had caused the mistaken perceptions.”

“Experiments” are directed toward measurement of a causal relationship between one or more variables by accounting for and eliminating error, “noise,” or other occurring (but not causal) elements, and then statistically removing them. A controlled experiment is “a procedure for testing cause-and-effect relationships within a setting that permits maximum control over extraneous variation and allows the experimenter to observe the effect of one variable on another in such a way as to demonstrate that no other variable could have produced the same effect.” The traditional controlled experiment utilizes a treatment group and a control group; only by comparing results from the two groups can the scientist understand the results. In a clinical trial to test the effects of a new drug therapy, for example, a placebo or other control mechanism is considered an essential component of the experimental design that allows the scientist to estimate the occurrence of false positives (i.e., positive responses not attributable to drug therapy). FDA regulations require well-controlled studies to demonstrate drug effectiveness; such regulations have been upheld against challenges that the requirement improperly slows the approval process or otherwise is unduly burdensome. Of course, in Daubert cases regarding causation as established by a “hard” science, a well-developed body of jurisprudence exists to indicate that a proponent of expert evidence must demonstrate methodological requirements such as controls to isolate alternate causation. The Lanham Act courts can benefit from the analysis in these cases.

176. Id. (emphasis added).
179. Controls in this setting allow the scientist to attempt, as much as is possible, to account for different factors that may influence an outcome such as test subjects’ health, genetic background, age, gender, and habits.
180. See 21 C.F.R. § 314.510.
182. In re Breast Implant Litig., 11 F. Supp. 2d 1217, 1228 (D. Colo. 1998) (finding that “case reports are not reliable scientific evidence of causation, because they simply describe
The problem with false positives in Lanham Act cases is illustrated by the Second Circuit decision in Schering v. Pfizer. There, the appellate court appeared to assume that if doctors told surveyors that “nonsedating” was a message conveyed—or impression left—by a ZYRTEC sales representative, it must be the case that such responses were “caused” by a false and misleading promotional statement. The Green survey, which provides the control feature that was missing from Schering’s initial evidence, dramatically reveals the fallacy in that assumption. The court’s confusion is betrayed by the repeated emphasis that the “surveys tended to corroborate one another.” Pfizer’s survey expert, Dr. Ivan Ross, attempted to explain that the “consistent” survey results (which were never higher than 20 percent) could simply reflect a consistent noise level; that is, a consistent measurement of responses that are not causally linked to the promotional statement conveyed by the ZYRTEC sales representative. “Noise,” which includes preexisting views of respondents and reactions to extrinsic sources of information, is especially likely when, as in this case, the survey population is sophisticated (physicians) and the product surveyed is well known (ZYRTEC).

reported phenomena without comparison to the rate at which the phenomena occur in the general population or in a defined control group. . . . [T]hey do not isolate and exclude potentially alternative causes . . . and do not investigate or explain the mechanism of causation.”); Haggerty v. Upjohn Co., 950 F. Supp. 1160, 1164 (S.D. Fla. 1996), aff’d, 158 F.3d 588 (11th Cir. 1998) (“[T]he generally accepted view in the scientific community is that [the expert’s] methodology [case reports, spontaneous reports of adverse medical events collected by the FDA, and animal studies] can be used to generate hypotheses about causation, but not causation conclusions. . . . [S]cientifically valid cause and effect determinations depend on controlled clinical trials and epidemiological studies.”); Jones v. U.S., 933 F. Supp. 894, 898 (N.D. Cal. 1996) (“Plaintiffs’ evidence shows, at most, that there is anecdotal support for the hypothesis that an unexplained interaction between certain antibiotics and oral contraceptives may reduce the effectiveness of the contraceptives. However, Plaintiffs’ evidence fails far short of the proven, cause-and-effect relationship that is necessary to satisfy the Daubert standard.”), aff’d, 127 F.3d 1154 (9th Cir. 1997); Schmaltz v. Norfolk & W. Ry. Co., 878 F. Supp. 1119, 1122 (N.D. Ill. 1995) (“It is well settled that a causation opinion based solely on a temporal relationship is not derived from the scientific method and is therefore insufficient to satisfy the requirements of [Daubert].”).

183. 189 F.3d at 239-40.

184. As discussed in supra Part III, Professor Green showed a videotaped sales presentation to doctors that conveyed indisputably truthful FDA labeling data (e.g., that 13 percent of patients experience somnolence when taking ZYRTEC versus 6 percent who took placebo). The survey results showed a similar percentage of doctors as in the Schering surveys (~20 percent) who said that “nonsedating” was the message conveyed.

185. Schering’s survey expert, Mr. Bartolomeo, testified that a control was “impossible.” Bartolomeo Tr. 204:14-17.

186. 189 F.3d at 236; see also id. at 230 (stating that the fact the “surveys corroborate one another is certainly a point in their favor”).

187. The law recognizes that a sophisticated audience is less likely to be misled than an untrained consumer audience. See Sandoz Pharms. Corp. v. Richardson-Vicks, Inc., 902 F.2d 222, 229-30 (3d Cir. 1990) (denying preliminary injunction where representations about cough syrup directed at audience of physicians); Quaker State Corp. v. Castrol, Inc.,
Designing a properly controlled deceptive advertising survey is not difficult. A relatively simple survey model would expose a “test group” to an advertisement alleged to contain an implied falsehood; separately, the surveyor would expose a “control group” to the same advertisement, but without the allegedly misleading statement. Both the test group and control group would be asked the same questions, funneling down from open-ended questions to more closed-ended questions. If, hypothetically, 50 percent of the test group play back a false belief, but only 20 percent of the control group play back a false belief, it could be argued that the true level of deception is 30 percent (the difference between the positives of both groups). Stated differently, 20 percent of this pool of respondents would convey the false belief because of some influence other than the stimulus (whether it be simply error, such as inattention to the question or mistake, or whether the response reflects an extrinsic belief that cannot be causally connected to the defendant’s advertisement). If, in the same example, 50 percent of both the test group and control group report the false belief, then the offending statement cannot be said to be the cause of these survey responses.\(^{188}\)

In the past, relatively few courts discussed the importance of controls in deceptive advertising cases. However, many courts are beginning to understand their vital significance. Professor Jacob Jacoby notes:

Perhaps nowhere is this emerging sophistication [in survey law] as important as in the increasing recognition of the need for surveys to incorporate adequate “controls.” The language of the Lanham Act proscribes advertising that is likely to cause deception. Yet there can be no trustworthy or valid assessment of cause and effect unless surveys are intertwined with proper experimental designs (which, of necessity, involve the utilization of proper controls). A decade ago, very few surveys offered as evidence incorporated “controls” and courts appeared not to recognize the scientific necessity for and critical significance of this omission. This is hardly the case today.\(^{189}\)

Fortunately, notwithstanding cases such as Schering, a growing number of courts have suggested that controls are an

\(^{188}\) See, e.g., Diamond, Reference Guide on Survey Research, supra n.132, at 257-58.

\(^{189}\) Jacob Jacoby et al., Survey Evidence in Deceptive Advertising Cases Under the Lanham Act: An Historical Review of Comments from the Bench, 954 PLI/Corp. 83, 89 (1996).
“indispensable” design element of deceptive advertising surveys. In a leading Second Circuit case, the defendant’s commercial emphasized that its antacid, TUMS, contained calcium, whereas plaintiff’s competing antacid, MYLANTA, contained aluminum salt. While this advertisement was literally truthful, plaintiff’s theory was that the defendant’s reference to aluminum was a veiled attempt to imply, falsely, that MYLANTA antacid was harmful. In particular, plaintiff alleged that the defendant was exploiting a “popularly held, yet unsubstantiated concern that aluminum is associated with Alzheimer’s” disease. Plaintiff conducted a survey that exposed respondents to the defendant’s advertisement and then asked “What if anything does the commercial communicate to you about the aluminum . . . in MYLANTA?” Forty-five percent of respondents said “harmful” or words to that effect. The defendant’s expert, Dr. Wind, testified that the survey was fatally flawed for its failure to include a control group. The Second Circuit affirmed the denial of injunctive relief and held that a control, which could have identified “the portion of the survey population that held extrinsic beliefs prior to viewing an advertisement,” was an

190. 960 F.2d 294, 297 (2d Cir. 1992).
191. Id. at 299.
192. Id. at 300.
    Dr. Wind criticized the Bruno & Ridgway study for not taking into account information about calcium and aluminum already known to the respondents. Dr. Wind testified that a survey must control for information people already know about these ingredients in order to measure accurately which messages about them were communicated by the commercial. He testified that another group should be asked similar questions without having been shown “Ingredients—Revised.” He found the lack of such a “control group” to be a basic flaw in the Bruno & Ridgway study.
194. Id.; see also Volkswagen Aktiengesellschaft v. Uptown Motors, No. 91 Civ. 3347, 1995 WL 605605, at *2-4 (S.D.N.Y. May 11, 1995) (defining and discussing the importance of controls); Reed-Union Corp. v. Turtle Wax, Inc., 869 F. Supp. 1304, 1311 (N.D. Ill. 1994) (noting that control is “essential” to discern survey responses which exist in the market independent of the alleged conduct), aff’d, 77 F.3d 909 (7th Cir. 1996); Major League Baseball Props., Inc. v. Sed Non Olet Denarius, Ltd., 817 F. Supp. 1103, 1123 (S.D.N.Y 1993) (finding that without controls, one cannot tell what respondent is reacting to; hence, “the interpretation of [the survey] data is impossible, and any conclusions drawn from their data must be seen as meaningless”), vacated by settlement, 859 F. Supp. 80 (S.D.N.Y. 1994); Pfizer Inc. v. Miles, Inc., 868 F. Supp. 437, 447 (D. Conn. 1994) (giving no credit to a survey that “lacked experimental control”).
Trademark cases are in accord. See Simon Prop. Group L.P. v. mySimon, Inc., 104 F. Supp. 2d 1033, 1045 (S.D. Ind. 2000) (holding that a survey “fails to include adequate controls. . . . By failing to use such controls, SPG’s proposed survey amounts to little more than a meaningless word association or memory exercise”); Nat’l Football League Props., Inc. v. ProStyle, Inc., 57 F. Supp. 2d 665, 667-68 (E.D. Wis. 1999) (survey discredited for failure to use a control pursuant to Kumho Tire and Daubert).
“indispensable” component of the survey design. Without a control group, there is no way to know whether the respondents thought that aluminum was harmful because of the advertisement or because of some totally unrelated influence.

A similar result was reached in American Home Products Corp. v. Procter & Gamble Co., in which the court rejected the plaintiff’s survey for want of an adequate control even though the survey used control questions to take account of inattention. Inattention, however, is only one element of “noise,” a point that many courts overlook. Although beneficial, the survey’s control questions did not make up for the lack of a control group to account for preconceptions:

It is clear that in a false advertising action survey results must be filtered via an adequate control mechanism to screen out those participants who took away no message from the advertisement as well as to account for those consumers who may have brought to the survey certain publicly held preconceptions regarding the product. In fact, the Second Circuit has held, in Johnson & Johnson ... that where a portion of the survey population may have held extrinsic beliefs prior to viewing an advertisement, a control mechanism would likely be “indispensable.” In this case AHP’s expert has conceded that the survey’s built-in control mechanism could not in fact control for any preconceptions regarding analgesics that the survey population may have possessed, even though some type of “noise” filter was clearly demanded. The existence of OTC analgesic noise was confirmed by the fact that when the same survey was conducted substituting an ADVIL commercial for the ALEVE spot, virtually the same results were obtained. Absent a valid noise filter, the Court is unable to discern if AHP’s survey results are attributable to the advertisement or are instead attributable to consumers being bombarded with years of OTC advertising from which their pre- or misconceptions of these products have developed.

196. Id.
198. Many courts equate “noise” with mistake or inattention. See, e.g., Ironclad, L.P. v. Poly-America, Inc., No. Civ. A. 3:98-CV-2600, 2000 WL 1400762, at *8 (N.D. Tex. July 28, 2000) (“The most significant challenge to the survey evidence is to the lack of a control. There is often general background noise in survey figures representing those people who are ‘bored, hurried, or just plain contrary.’” J. McCarthy, Trademarks and Unfair Competition § 32:54, at 784. Many courts have required control questions in order to filter out these responses.”). In fact, “noise” also includes responsive answers, i.e., answers on point from people paying attention to the question. The problem is that the answer was influenced by something other than the advertising that is being tested.
Accordingly, the Court must discount plaintiff’s television commercial survey.199

The utility of controls in deceptive advertising cases also is reflected by those cases in which controls have been found to compensate for other shortcomings in the survey design. For example, in another case involving antacids, Novartis Consumer Health, Inc. v. Johnson & Johnson*Merck Consumer Pharmaceuticals Co.200 the court granted plaintiff’s motion for preliminary injunction to prevent Johnson & Johnson from using the designation “Nighttime Strength” to describe its MYLANTA antacid products on the theory that, although the antacids provide relief for only 30 to 60 minutes, the tagline misleads consumers into believing that it will provide relief through the night.201 Plaintiff constructed a survey that was divided into two cells, an experimental cell and a control cell. One cell was exposed to the label “Mylanta Nighttime Strength,” and the control cell was exposed to the label “Extra Strength Mylanta.”202 The court found that the control cell “functions as a baseline and provides a measure of the degree to which respondents are likely to give an answer indicating a period of time of relief, not as a result of the name or labeling, but because of other factors, such as the survey’s questions, the survey’s procedures, the nature of the product, or some other potential influence on a respondent’s answer such as preexisting beliefs.”203 The court believed that inclusion of a control group ameliorated other potential defects in survey design. For example, in discussing defendant’s claim that the survey included leading questions, the court held that the use of a control group cured the error: “[B]ias that may have resulted from any leading questions was eliminated by use of the control group.”204 A control group, therefore, may be seen as a core component of a survey design that can go a long way toward curing other survey defects.

C. Danger of Using Surveys as Memory Tests to Prove Literal Falsehoods

In the Schering v. Pfizer case, the issue of controls was strenuously argued but never seriously considered—a fact due to the holding by the Second Circuit that Schering’s surveys could be admitted to prove “literal falsehoods,” that is, as evidence of doctors’ memories of the precise words actually spoken by the sales

---

201. Id. at 364.
202. Id. at 365.
203. Id. at 365 n.10.
204. Id. at 365.
representative. This theory of literal falsehood drove the final result in Schering, yet the reliability of surveys to show memories of precise words spoken by a sales representative merits scrutiny.

On remand, the district court agreed that controls are required for implied falsehood cases (citing the Johnson & Johnson case) but concluded that a control was not necessary in the instant case because Schering used the surveys to prove literal falsehoods.\(^{205}\) The trial court obviously was constrained by the Second Circuit to abandon its prior belief that permitting “memory” or “recall” surveys to show literal falsehoods would lead to a “virtual destruction of the hearsay rule.”\(^{206}\) The remand decision reflects the assumption of the Court of Appeals that surveys of physicians could reliably and accurately recreate the precise words that the Pfizer detail representatives spoke during the detail meeting (as opposed to a paraphrase by the physician).\(^{207}\) The Green survey experiment discussed earlier exposed the error of this assumption by showing that 25 percent of physicians told a surveyor that “nonsedating” was a main message of a ZYRTEC detail even though the video presentation never used this phrase.\(^{208}\) The result of the remand decision also reveals the confusion of the Second Circuit’s opinion in Schering. The Second Circuit devoted substantial analysis to Schering’s theory of implied falsehood\(^{209}\) (a theory that the district court on remand denied was proffered).\(^{210}\) The Second Circuit then held, without much analysis, that Schering should be free to pursue a theory that the surveys can be used “to establish literal falsehoods under the residual hearsay rule” as “memory surveys offered to prove the facts remembered.”\(^{211}\) This holding is not only unsupported by the analysis in the opinion, it is unprecedented. Surveys in deceptive

---

208. See discussion of Green study in supra Part III.
209. Schering, 189 F.3d at 228-30.
210. Schering, 2000 WL 718449, at *10 n.15:
The parties have spilled a great deal of ink over the question whether Schering has or has not asserted a claim based on implied falsehood. The Court does not reach this question. An implied falsity claim—that “while the advertisement is literally true it is nevertheless likely to mislead or confuse consumers,” Johnson & Johnson*Merck Consumer Pharm. Co. v. Smithkline Beecham Corp., 960 F.2d 294, 297 (2d Cir. 1992)—requires that a plaintiff show not only the residual belief into which the consumer was misled, but what the literally true message was that misled the consumer. The Court finds that there is not (or, at least, there has not been pointed out to the Court) sufficient evidence from which to decide such a claim, if Schering has asserted it. The surveys do not seem to have been designed to elicit both of the necessary facts, and any findings in this regard would be speculation. The Court notes that the calling of at least some physicians would be particularly useful in deciding an implied falsehood claim.
211. 189 F.3d at 221.
advertising cases are used to show implied falsity; surveys have not been used to show the content of the promotional message itself.

Is a survey of physicians to establish the precise words told to the physician by a sales representative “reliable”? An examination of the history and policy rationales undergirding the hearsay rule is beyond the scope of this article. Suffice it to say, throughout American jurisprudence, courts have echoed the United States Supreme Court in 1813: “Its intrinsic weakness, its incompetency to satisfy the mind of the existence of the fact, and the frauds which might be practiced under its cover, combine to support the rule that hearsay evidence is totally inadmissible.”

Far from satisfying the Supreme Court’s command in 1999 that all technical evidence must be relevant and reliable, surveys consisting of hearsay recollections of remembered facts have all of the indicia of unreliability:

The exclusion of it, in other cases, stands upon the general consideration that it is not upon oath; that the party affected by it has no opportunity of cross-examination; that it often supposes better evidence behind; that it is peculiarly liable to be obtained by fraudulent contrivances; and above all, that it is exceedingly infirm, unsatisfactory, and intrinsically weak in its very nature and character. On these accounts judges in modern times have leaned against any extension of it, as being subversive of the security of the titles of parties to property.

Hearsay obviously has high potential for unreliability due to “faults in the perception, memory, or narration of the declarant.” As anyone who has played the children’s “telephone game” knows, a message always becomes distorted to some degree through verbal transmission. Factual information rarely is heard or remembered

---

213. Kumho Tire, 526 U.S. at 141.
215. 5 Weinstein’s Fed. Evid., § 802[3] (2d ed. 1997). See T. Harris Young & Assoc., Inc. v. Marquette Elec., Inc., 931 F.2d 816, 827-28 (11th Cir. 1991) (excluding a telephone survey of hospital employees asking what the defendant’s representatives said to the hospital employee about the use of defendant’s paper in EKG and stress testing machines): [Plaintiff] admits that it was trying to prove that Marquette employees had made tortious comments. The relevant issue is thus whether Marquette employees made those comments, not the state of mind of the hospital employees. Rule 803(3) is also of no avail to [Plaintiff] because [Plaintiff] tried to use the alleged statements “of memory or belief” by the hospital employees “to prove the fact remembered or believed.” That is, [Plaintiff] tried to use the out of court statements that “Marquette said . . .” to prove that “Marquette said . . .” The prohibition against such use of statements “is necessary to avoid the virtual destruction of the hearsay rule.” Fed. R. Evid. 803, notes of Advisory Committee on 1972 proposed rules.
with accuracy. The listener often remembers portions of a conversation or its gist; however, the listener fills in "gaps" in such memory based upon logical inferences of what the listener thinks "probably happened." The listener's "biases, expectations, and past knowledge are all used in the filling-in process, leading to distortions in what [the listener] remember[s]."

Recent research confirms the substantial imperfections of human memory in recounting words previously spoken by another person. Experiments show that the memory flaws known as "misattribution" and "false recognition" can occur "when we fail to recollect specific details of an experience, and at the same time recall the general sense of what happened." For example, when laboratory subjects were read a list of semantically associated words including candy, sugar, sour, bitter, and taste, many respondents incorrectly recalled that they previously heard the word "sweet."
Lanham Act surveys should be designed to demonstrate a causal link between the advertiser’s message and confusion or deception in a projectable sample of the relevant consumer audience. This cannot be accomplished with a “memory test.” Indeed, a few months before deciding Schering, a different panel of the Second Circuit affirmed the exclusion of a survey on the basis that “the survey was little more than a memory test, testing the ability of the participants to remember the names of the shoes they had just been shown and gave no indication of whether there was a likelihood of confusion in the marketplace.”

Given the imperfections of human memory, reliability of recall also diminishes as the time between exposure to the stimulus and participation in the survey lengthens. To the extent memory surveys are to be used in a litigation context, questioning must follow close on the heels of the relevant event. Even with “perception” surveys, courts have criticized surveys where even a brief time interval separated the survey response from the respondent’s exposure to the promotional message. Any “interval” undermines the position that such perceptions are “present sense” perceptions, admissible as an exception to the hearsay rule. In Schering, the surveys were conducted one day to one week after the detail session, which was yet another reason to worry about the surveys’ reliability.

memory test, subjects decide whether each of several words had been read aloud earlier: sewing, door, needle, sleep, candy, awake. Most of the time, people correctly remember that they had earlier heard sewing and awake, and correctly state that they had not heard door and candy. More interestingly, people frequently claimed—confidently but incorrectly—that they heard needle and sleep. You might even have made this error yourself as you looked over the test words.

Id. at 98.

221. Starter Corp. v. Converse, Inc., 170 F.3d 286, 297 (2d Cir. 1999). The Second Circuit should consider the potential inconsistency between Schering and Starter together with the impact of the 1999 Supreme Court decision in Kumho Tire. In defense of the Second Circuit, Schering, Starter, and Kumho Tire all were decided in 1999 before the possible connection between these cases may have been apparent.


223. Pittsburgh Press Club v. United States, 579 F.2d 751, 759 (3d Cir. 1978) (rejecting survey in part because respondents “were not being asked for a present impression” but rather to remember events in the past) (emphasis added).

224. 189 F.3d at 236.

225. Methodological flaws in a survey undermine the ability of the proponent to show that the survey has the requisite guarantees of trustworthiness required by the residual exception to the hearsay rule (Rule 807), see C.A. May Marine Supply Co. v. Brunswick Corp., 649 F.2d 1049, 1052, 1055 (5th Cir. 1981); Pittsburgh Press Club, 579 F.2d at 756-60. The residual exception ultimately was used in Schering, by reason of the Second Circuit’s ruling, to admit the surveys to prove facts remembered. See Schering, 2000 WL 718449, *4
The only case other than Schering that has permitted survey data to shed light on the content of pharmaceutical sales representatives' verbal statements is Zeneca, Inc. v. Eli Lilly and Co.\textsuperscript{226} Zeneca does not support the proposition that such surveys can be used as dispositive evidence in deceptive advertising cases, however.\textsuperscript{227} Zeneca concerned Eli Lilly’s drug product, EVISTA. Zeneca alleged that Eli Lilly was promoting EVISTA not only for the prevention of osteoporosis (which is an approved indication of EVISTA) but also for the prevention of breast cancer (which was not an approved indication).\textsuperscript{228} Zeneca’s surveys, which asked physicians to recall the message conveyed by Eli Lilly sales representatives during detail sessions, indicated that some of the Eli Lilly representatives communicated that EVISTA was effective for the prevention of breast cancer. However, unlike the Schering case, Zeneca did not offer survey evidence as dispositive proof of the issue in dispute. To the contrary, Zeneca introduced the survey evidence only as corroboration for Zeneca’s substantial direct evidence. In particular, Zeneca introduced written notes by sales representatives summarizing topics discussed, training scripts for sales representatives, eyewitness testimony, and Eli Lilly executive testimony, all of which showed that Eli Lilly was instructing its sales representatives to promote EVISTA as a treatment for reducing the risk of breast cancer.\textsuperscript{229}

Furthermore, the surveys in Zeneca were less susceptible to ambiguity than were the surveys in Schering. Doctors were unlikely to confuse “osteoporosis” and “breast cancer” when responding to the Zeneca surveys. In the Schering case, by contrast, because of the subtlety and ambiguity of the terminology, one could not tell whether the survey responses reflected a prohibited message (e.g., ZYRTEC is “essentially nonsedating.”) or


\textsuperscript{227} Cf. Harold P. Weinberger & Jonathan M. Wagner, On the Advertising Battlefront: Fiercest Lanham Act Conflicts Waged By Large Pharmaceutical Companies Over Reps’ Oral Claims, 224 N.Y.L.J. 553 (July 24, 2000) (arguing that “Schering and Zeneca provide a clear and viable road map” and provide a method “for a plaintiff to prove that a competitor’s sales reps are orally making false claims.”).

\textsuperscript{228} 1999 WL 509471, at *1.

\textsuperscript{229} Id. at *9. The court’s treatment of another survey in Zeneca is instructive. The survey concluded that 11 percent of the physicians surveyed wrote prescriptions for Evista for breast cancer prevention alone, and that 35 percent had written at least some of their prescriptions primarily for breast cancer prevention. Id. at *30. The court is careful to note, “This study does not, however, indicate that these prescribing patterns of physicians are the result of statement[s] made by Eli Lilly, although the study is evidence that Evista competes with tamoxifen in the marketplace.” Id.
indisputably truthful labeling data (e.g., ZYRTEC causes 14 percent of patients to experience somnolence compared to 6 percent taking placebo).

It is worth noting that, when dealing in the context of advertising claims made verbally by sales representatives, a plaintiff also must ensure that proof offered establishes more than just isolated or “off the cuff” utterances, which typically are not actionable under the Lanham Act. Instead, the plaintiff must show that the alleged deception is “disseminated sufficiently to the relevant purchasing public to constitute advertising or promotion within that industry.” Without such proof, it is difficult to invoke the extraordinary remedy of a preliminary injunction. As one court analyzed the issue, “isolated individualized oral comments about competitors are at the opposite pole from clearly definable media advertising or printed material containing specific verifiable or disprovable statements and given wide distribution in commerce. In the latter situation, both certainty concerning what is said (if not always concerning its truth) and the strength of the probable impact are more readily established or foreseen, permitting interim relief.”

As demonstrated by the survey experiment conducted by Professor Wind, physicians do not uniformly use the term “nonsedating” in the same way that Schering posited (as causing no statistically significant incidence of “sedation”). To the contrary, many physicians characterize ZYRTEC as a “nonsedating antihistamine” while fully recognizing that ZYRTEC causes a statistically higher incidence of somnolence than does placebo or CLARITIN. The court in the Schering case never

---

230. Oral statements by a company’s sales representative can constitute “commercial advertising or promotion” under the Lanham Act. In addition to Schering and Zeneca, see also Avon Prods., Inc. v. S.C. Johnson & Son, Inc., 984 F. Supp. 768, 772, 796 (S.D.N.Y. 1997).

231. See Boule v. Hutton, 70 F. Supp. 2d 378, 389 (S.D.N.Y. 1999) (holding that commercial advertising and promotion connote an affirmative “reaching out to consumers” in a “proactive” communication; hence “isolated” and “reactive” statements—uttered only to three people—could not be considered “commercial advertising or promotion”). In the case of verbal statements made by pharmaceutical representatives, courts should not permit a Lanham Act case to go forward where the allegation at best is one of a “rogue rep” problem—in other words one or two sales representatives make wanton, ultra vires, or spur of the moment statements inconsistent with the company’s marketing campaign as a whole.


233. Licata & Co., Inc. v. Goldberg, 812 F. Supp. 403, 408 (S.D.N.Y. 1993). The court in Licata also noted that it may be more difficult to establish that verbal statements are material. See id. (“The Lanham Act’s antideception provision . . . would be trivialized if it were applied to statements in oral conversations by an individual sales representative to an individual customer concerning matters which an ordinary listener would recognize as personal opinion as opposed to representations of hard definable facts, such as product descriptions.”).

234. See supra Part III.
inquired or asked for evidence about the way practicing physicians who prescribe antihistamines use terminology such as “sedation,” “nonsedating,” “somnolence,” “drowsiness,” “sluggishness,” “fatigue,” and related concepts, even though the “special knowledge” of a “sophisticated audience” is highly relevant in false advertising cases. As Alice discovered in Through the Looking Glass, it is important to understand how the speaker defines his terminology. The FDA-approved labeling uses the term “somnolence,” which connotes a milder symptom than does “sedation,” a term that conjures something on par with a tranquilizer or anesthetic. As learned intermediaries, physicians are expected to critically evaluate scientific and other factual information, integrate it with their own experience and learning, and reach conclusions that they articulate in their own words. Schering’s survey data itself demonstrated that some doctors would call ZYRTEC “nonsedating” yet still understand that ZYRTEC causes drowsiness. Physicians were more likely to characterize the data in a way that is consistent with their understanding and experience rather than playback as would a tape recorder the precise words used by the sales representative. Because the survey responses were ambiguous, the failure to use follow-up questions or funneling techniques to clarify responses is itself a significant flaw that indicates the surveys

235. Cf. Cumberland Packing Corp. v. Monsanto Co., 32 F. Supp. 2d 561, 582-83 (E.D.N.Y. 1999) (rejecting plaintiff’s implied falsehood claim that name “NutraSweet” misleadingly implies that the product is a sweetening ingredient and not a sugar substitute on the ground that plaintiff failed to present any evidence to show that the target consumers—as opposed to the parties themselves—made a meaningful distinction between the terms). 236. See Utah Med. Prods., Inc. v. Clinical Innovations Assocs. Inc., 79 F. Supp. 2d 1290, 1315-16 (D. Utah 1999), aff’d, 251 F.3d 171 (Fed. Cir. 2000) (in suit challenging defendant’s advertising of its intrauterine catheter as literally false because advertisements claimed the catheter was “sensor-tipped” when the transducer was not actually located at the tip of the catheter, the court excluded plaintiff’s expert under Daubert and Kumho Tire for failure to analyze “how [the advertisements] were perceived among the clinicians whom the advertisements targeted” in light of their special knowledge and training). 237. See supra n.1. 238. Some physicians obviously use the term “nonsedating” not to indicate literally “no sedation” but rather as a short-hand way to lump or classify ZYRTEC along with ALLEGRA and CLARITIN as part of the “new” generation of antihistamines. The failure to discern physician’s terminology and usage may have resulted in the collection of survey data that was irrelevant to the litigation issue. Determining whether a survey is relevant to answering the litigation is highly contextual. See Sterling Drug, Inc. v. Bayer AG, 14 F.3d 733, 741 (2d Cir. 1994); Continental Plastic Containers, Inc. v. Owens-Brockway Plastic Prods., Inc., No. 95 C 4670, 1996 WL 284960, at *3 (N.D. Ill. May 28, 1996) (survey “irrelevant” to litigation question). 239. Cf. Wash. Legal Found. v. Friedman, 13 F. Supp. 2d 51, 70 (D.D.C. 1998), vacated in part on other grounds, 202 F.3d 331 (2000).
were entitled to little or no weight. More importantly, physicians’ usage of the terminology indicates that they may internally process or remember the statistical somnolence data conveyed by the Pfizer representative in more general terms. The physicians remember the “gist” of what was said, encapsulating the message in terminology that may have clinical or other significance to them. This informs the issue of whether memory tests possess sufficient indicia of trustworthiness to admit them into evidence under the “rarely” used residual hearsay exception to support the “drastic” remedy of a preliminary injunction.

V. CONSTITUTIONAL AND PUBLIC POLICY CONSIDERATIONS

Holding plaintiffs to their burden in deceptive advertising cases obviously fulfills the federal court’s gatekeeping function and obligation to insist that plaintiffs shoulder their legal and scientific burdens. Allowing deceptive advertising claims to proceed based on shaky scientific evidence may chill informative speech that is constitutionally protected and offers public benefits in fostering informed consumer decisionmaking. The Supreme Court has


241. Schering, 189 F.3d at 232.


243. Lanham Act litigation may not be used to intrude upon constitutionally protected commercial speech: that is, truthful and nonmisleading advertising. Thus, a Lanham Act plaintiff’s obligation to demonstrate that the challenged advertising is “false or misleading” is more than a statutory requirement. If a Lanham Act plaintiff cannot meet this burden, then the statement at issue is truthful, nonmisleading commercial speech that not only is outside the ambit of the Lanham Act, but is constitutionally protected. See, e.g., Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council, Inc., 425 U.S. 748 (1976); see also Cent. Hudson Gas & Elec. Corp. v. Pub. Serv. Comm’n, 447 U.S. 557, 561-62 (1980) (holding that “[c]ommercial expression not only serves the economic interest of the speaker but also assists consumers and furthers the societal interest in the fullest possible dissemination of information”). Although not protected with the same intensity as “core” speech, see, e.g., R.A.V. v. City of St. Paul, 505 U.S. 377, 422 (1992) (Stevens, J., concurring) (“Our First Amendment decisions have created a rough hierarchy in the constitutional protection of speech. Core political speech occupies the highest, most protected position; commercial speech and nonobscene, sexually explicit speech are regarded as a sort of second-class expression; obscenity and fighting words receive the least protection of all.”), truthful, nonmisleading commercial speech enjoys constitutional protection because of its substantial public value in contributing to the marketplace of ideas and facilitating efficient consumer decisionmaking. Bates v. State Bar of Arizona, 433 U.S. 350 (1977); see also Rubin v. Coors Brewing Co., 514 U.S. 476, 481-82 (1995); Ibanez v. Fla. Dep’t of Bus. & Prof’l Regulation, 512 U.S. 136, 142 (1994) (finding that “because disclosure of truthful, relevant information is more likely to make a positive contribution to decisionmaking than is concealment of such
held that “the consumer’s concern for the free flow of commercial speech often may be far keener than his concern for urgent political dialogue.” Failure to enforce plaintiffs’ burdens in deceptive advertising cases may also indirectly encourage “puffing” advertising, not actionable under the Lanham Act, which attempts to sell images rather than to convey the type of factual data about the “availability, nature, and prices of products and services” that fosters informed consumer decisions.

The Lanham Act does not erect formidable barriers to sue: any rival has standing to sue; a plaintiff has no obligation to show harm to consumers or scienter on the part of the defendant; and a plaintiff can obtain emergency injunctive relief within a matter of days or weeks. Now the Second Circuit in Schering has suggested that survey evidence may be presumptively admissible, with flaws going to weight rather than admissibility. Additionally, rivals may be naturally motivated to use false advertising litigation as another method of competition. Given these factors, adherence to

244. Bates, 433 U.S at 364:

The listener’s interest is substantial: the consumer’s concern for the free flow of commercial speech often may be far keener than his concern for urgent political dialogue. Moreover, significant societal interests are served by such speech. Advertising, though entirely commercial, may often carry information of import to significant issues of the day. And commercial speech serves to inform the public of the availability, nature, and prices of products and services, and thus performs an indispensable role in the allocation of resources in a free enterprise system. In short, such speech serves individual and societal interests in assuring informed and reliable decisionmaking. See also Edenfield v. Fane, 507 U.S. 761, 767 (1993) (holding that a Florida ban on in-person solicitation by certified public accountants “threatens societal interests in broad access to complete and accurate commercial information that the First Amendment is designed to safeguard”). The Court in Edenfield commented that a “seller has a strong financial incentive to educate the market and stimulate demand for his product or service, so solicitation produces more personal interchange between buyer and seller than would occur if only buyers were permitted to initiate contact.” The Court also said that the “personal interchange” between a potential buyer and seller carries “benefits [that] are significant” in that it allows the buyer and seller to discuss the product, the buyer’s needs, and the way in which the product compares to alternatives in the marketplace. Id.; cf. Metromedia, Inc. v. City of San Diego, 453 U.S. 490, 497 (1981) (holding that outdoor billboard advertising “produces numerous direct and indirect benefits to the public. Valuable commercial, political and social information is communicated to the public through the use of outdoor advertising. Many businesses and politicians and other persons rely upon outdoor advertising because other forms of advertising are insufficient, inappropriate and prohibitively expensive”); see also Bolger v. Youngs Drug Prods. Corp., 463 U.S. 60, 67-68 (1983) (voiding federal statute prohibiting unsolicited mailing of contraceptive advertisements as an undue burden on the First Amendment; and holding that the mailings “constitute commercial speech notwithstanding the fact that they contain discussions of important public issues such as venereal disease and family planning”). Hence, the Court in Bolger held that the speech at issue was “commercial” and, therefore, not entitled to the more vigorous protection afforded to political or “core” speech; nevertheless, the statute failed to survive the Central Hudson four-part analysis.

the federal courts’ obligation to insist upon relevant and reliable scientific evidence becomes all the more important from a public policy perspective.

Some of the information exchanged in this commercial marketplace is “vital”; other information has “slight” value or may even be regarded in some quarters as unsavory or immoral. However, the vitality of the marketplace of ideas depends on broad content-neutral protection. For example, paternalistic notions that consumers should be kept ignorant of certain information for their own protection has consistently been rejected as incompatible with important First Amendment values. As one court has explained:

Robust debate between competitors on matters of opinion, and claims that one product or service is far superior to that of rivals, are encouraged as part of the hurly-burly inherent in a free market system, and indeed an open society. The deceptive potentialities of alleged misstatements balanced against the restrictive impact of prior restraint on freedom of commercial speech must be evaluated—especially on an application for preliminary restraint to be imposed prior to trial—based on the probable impact on those to whom the speech is directed.

Thus, the Lanham Act and First Amendment abut each other; the implications of failing to hold Lanham Act plaintiffs to their burden has constitutional significance. First Amendment considerations pervade deceptive advertising litigation, even in cases where a violation of the statute can be established. For example, any judicial order requiring corrective advertising constitutes state-compelled speech that must satisfy constitutional scrutiny; that is, the order must directly advance a “substantial” government interest and must not be more extensive than is necessary to serve that interest. If plaintiffs are permitted to

246. Edenfield, 507 U.S. at 767.
247. 44 LiquorMart, Inc. v. Rhode Island, 517 U.S. 484, 497 (1996) (stating that “paternalistic assumption that the public will use truthful, nonmisleading commercial information unwisely cannot justify a decision to suppress it”).
249. See Central Hudson, 447 U.S. at 564. Corrective advertising is a remedy that is more difficult to obtain and is disfavored where the proposed advertising “would tend to confuse or misinform the public as to defendant’s products.” Burndy Corp. v. Teledyne Indus., Inc., 584 F. Supp. 656, 670 (D. Conn.), aff’d, 748 F.2d 767 (2d Cir. 1984). To request corrective advertising, a party has to demonstrate that the defendant communicated a deceptive message, that at least a substantial portion of the target audience in fact was deceived, that there would be continuing deception even if the defendant ceased the promotional message, and the proposed corrective message would not itself lead to confusion or deception. See Warner-Lambert Co. v. FTC, 562 F.2d 749, 762 (D.C. Cir. 1977); cf. Peel v. Attorney Registration & Disciplinary Comm’n of Ill., 496 U.S. 91, 108 (1990). If plaintiff carries this burden, corrective advertising may be needed to dispel lingering deception from prior false advertising. See Energy Four, Inc. v. Dornier Medical Systems, Inc., 765 F. Supp.
obtain relief on deceptive advertising claims grounded upon survey or other evidence that lacks credibility and trustworthiness, the flow of truthful and accurate commercial information will be interrupted, causing a public harm. The rationale for affording constitutional protection to truthful and accurate commercial speech is not inconsistent with the goals of the Lanham Act in providing a remedy for deceptive commercial speech. Both forms of public law are aimed at ensuring an efficient marketplace of ideas providing the maximum flow of accurate information to consumers, that is, a vibrant marketplace free from substantial defect.250

VI. CONCLUSION

An analysis of the reliability and relevance of survey evidence in deceptive advertising cases must be conducted in light of the statutory and constitutional restrictions and scientific requirements applicable to the survey proponent’s case. In implied falsehood cases, relief is available when a “reliable” (methodologically sound) and “relevant” survey can both (1) measure a false belief in a substantial segment of a target audience; and (2) show that such false belief is causally linked to an advertiser’s promotional statement. Deceptive advertising surveys must be designed not to measure generalized consumer “attitudes”; instead, the survey must control for “noise” and other extraneous factors in order to discern the actual cause-and-effect relationship, if any, between an advertiser’s statement and a deceptive message. As such, these surveys, like all cause-and-effect scientific experiments, must be designed within the established principles of the scientific method, which include appropriate control groups and control questions.

Additionally, in express falsehood cases involving sales representatives who make in-person sales calls, courts should be wary of any attempt to use “surveys” as “recall tests” or “memory tests” to ascertain the precise content of promotional messages delivered verbally. There may be occasions, such as in the Zeneca case, where such recall data may provide useful corroborating information, and when the issue being tested is obvious and memorable and not “precise yet subtle.” Ordinarily, however, such recall tests should be recognized as intrinsically vulnerable to the


250. An analogous goal is served by excluding from First Amendment protection certain narrow categories of utterances (e.g., threats or calls to imminent violence) because they tend to impair the workings of the very marketplace of ideas itself. See Chaplinsky v. New Hampshire, 315 U.S. 568, 572 (1942) (categorizing as unprotected by the First Amendment words which “by their very utterance inflict injury or tend to incite an immediate breach of the peace”).
imperfections of human memory and should therefore be excluded or discounted accordingly.

When considering surveys in deceptive advertising cases, federal courts should approach their gatekeeping obligations with substantially the same care, scrutiny, and methodology as courts follow when considering forensic chemistry testimony, medical testimony, and other traditional “scientific” evidence. Courts should bring to bear the relevant scientific principles to ensure that deceptive advertising suits are not prosecuted with unreliable surveys. By enforcing these legal and scientific requirements, federal courts will effectuate the goals of the Lanham Act and the First Amendment to ensure that deceptive advertising litigation facilitates, rather than undermines, the vibrancy of the commercial marketplace.